DOCUMENT RESUME

ED 033 469 EA 002 595

AUTHOR	Bressler, Marvin; Tumin, Melvin M.
TITLE	Evaluation of the Effectiveness of Educational Systems. Final Report. Volume I.
INSTITUTION	Princeton Univ., N.J.
Spons Agency	Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No	BR-6-2023
Pub Date	Apr 69
Grant	OEG-1-7-1062023-0164
Note	447p.
EDRS Price	EDRS Price MF-$1.75 HC Not Available from EDRS.
Descriptors	Classification, *Comparative Education, *Cross Cultural Studies, Developed Nations, *Educational Objectives, Educational Programs, Educational Theories, *Models, Performance Criteria, Primary Education, *Program Evaluation, Secondary Education, Social Systems, Taxonomy
Identifiers	England, France, Italy, Netherlands, United States, West Germany

Abstract

Volume I discusses the background, findings and analysis, and conclusions and recommendations of a two-volume study aimed at developing a cross-national model to enable the United States and five European democracies, England, France, Italy, the Netherlands, and West Germany, to make more adequate estimates of the effectiveness of their primary and secondary schools. Two major tasks attempted were the development of taxonomy for the classification of claimed and confirmed educational outcomes and a comprehensive measure to indicate a gross educational product. Appendices include an annotated bibliography on the aims and outcomes of education and a case study of the application of the model of educational goals and outcomes to materials published by educational sources in New Jersey. Two articles, "Evaluation of Educational Goals" and "Problems in the Theory and Method of the Classification of Outcomes," are included. A related document is Volume II, EA 002 596. [Not available in hard copy due to marginal legibility of original document.] (MF)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED033469

FINAL REPORT

Cooperative Research Project No. 6-2023

Contract No. OEG-1-7-1062023-0164

EVALUATION OF THE EFFECTIVENESS OF EDUCATIONAL SYSTEMS

Marvin Bressler and Melvin M. Tumin
in collaboration with
S. J. C. vanEyndhoven
E. Fitzner
Ettore Gelpi
Charles Geronomi
Johan Goudsblom
Joseph Majault
Bernd Stickelmann
and
William Taylor

and with contributions by
Iseli K. Krauss
Joseph Lennards
James H. McGregor
Anthony H. Smith
Jan S. Smith
and
Margaret M. Waldron

Princeton University
Princeton, New Jersey

In two volumes: Volume I
April 1969

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

Best Copy Available

EA 002 595

FINAL REPORT
Cooperative Research Project No. 6-2023
Contract No. OEG-1-7-1062023-0164

EVALUATION OF THE EFFECTIVENESS OF EDUCATIONAL SYSTEMS

Marvin Bressler and Melvin M. Tumin
in collaboration with
S. J. C. vanEyndhoven, E. Fitzner, Ettore Gelpi,
Charles Geronomi, Johan Goudsblom, Joseph Majault,
Bernd Stickelmann, and William Taylor
and with contributions by
Iseli K. Krauss, Joseph Lennards, James H. McGregor,
Anthony H. Smith, Jan S. Smith, and Margaret M. Waldron

Princeton University
Princeton, New Jersey

In two volumes: Volume I
April 1969

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TABLE OF CONTENTS

VOLUME I

A. INTRODUCTION, METHODS, SUMMARY

1. Background of the Study

B. FINDINGS AND ANALYSIS

2. Education as a Social System: Functions, Structures, and Strains

3. Defining the Goals of Education: Requirements and Ambiguities

4. Establishing the Claims of Education: Requirements and Ambiguities

5. The Evaluation of the Effectiveness of Educational Systems: Some Problems

6. Cross-National Comparisons: Some Problems and Solutions

7. Alternative Models of Educational Systems and Their Evaluation

James H. McGregor

C. CONCLUSIONS AND RECOMMENDATIONS

8. Toward Evaluation: Some Indispensable Elements in the Measurement of the Effectiveness of Educational Systems

D. APPENDICES

1. Aims and Outcomes of Education: A Selected Annotated Bibliography

Iseli K. Krauss

2. Evaluation of Educational Goals
Josepn Lennards

3. Problems in the Theory and Method of the Classification of Outcomes
Anthony H. Smith and Jan S. Smith

4. Case Study: The Application of the Bressler-Tumin Model of Educational Goals and Outcomes to Materials Published by the New Jersey Department of Education and the Princeton, New Jersey, Regional Schools
Iseli K. Krauss and Margaret M. Waldron

VOLUME II

5. Educational Goals and Outcomes in England
William Taylor

6. Major Changes in French Education in the Twentieth Century
Joseph Majault and Charles Geronomi

7. The Aims of Education in Western Germany: Their Manifestations and Social Determinants
Bernd Stickelmann

8. The Aims, Values, and Content of Education as Manifested in the Curriculum in West German Schools
E. Fitzner

9. Educational Goals in Italy
Ettore Gelpi

10. Educational Goals in the Netherlands
S. J. van Eyndhoven

11. On the Definition of Goals and the Classification of Outcomes
Johan Goudsblom

12. Educational Problems and the Situation of Educational Research in the Netherlands
from Foundation for Educational Research in the Netherlands

VOLUME I

A. INTRODUCTION, METHODS, SUMMARY

1. Background of the Study

All Western democratic societies, American and European, rely on their schools to promote social and individual welfare. The diversity of tasks assigned to the educational system in the United States may be illustrated by the influential list of the components of "quality education" developed by a committee of the State Board of Education of the Commonwealth of Pennsylvania. The set of goals which follow probably command at least the nominal allegiance of a sizable segment of the American educational community.

1. To help every child acquire the greatest possible understanding of himself and an appreciation of his worthiness as a member of society.
2. To help every child acquire understanding and appreciation of persons belonging to social, cultural, and ethnic groups different from his own.
3. To help every child acquire to the fullest extent possible for him mastery of the basic skills in the use of words and numbers.
4. To help every child acquire a positive attitude toward school and toward the learning process.
5. To help every child acquire the habits and attitudes associated with responsible citizenship.
6. To help every child acquire good health habits and an understanding of the conditions necessary for the maintenance of physical and emotional wellbeing.
7. To help give every child opportunity and encouragement to be creative in one or more fields of human endeavor.
8. To help every child understand the opportunities open to him for preparing himself for a productive life and to enable him to take full advantage of these opportunities.
9. To help every child to understand and appreciate as much as he can of human achievement in the natural sciences, the social sciences, the humanities, and the arts.
10. To help every child to prepare for a world of rapid change and unforeseeable demands in which continuing education throughout his adult life should be a normal expectation.
11. To provide every child with equal access to quality education, and to its social rewards, regardless of differences in race, creed, national origin or social class, or talent.[1]

This array of objectives is typical in range, if not necessarily in content, of numerous efforts to define the functions of the schools. It is odd, therefore, with so much investment of hope that we know so little about the precise nature of the interrelationships between stated aims and actual outcomes. Apparently we now lack both the intellectual apparatus -- i.e. the standards, theories, concepts, indicators, tests, and raw data -- and sufficient will that would permit us to distinguish a "success" from a "failure." As Henry Dyer, who is among the most astute and knowledgeable of all observers of schools, has said: "Pupils, teachers, administrators, and policy-makers rarely give any though to the question of why they are going through all the ritualistic motions they think of as education." He also adds that the "overwhelming problems of survival in a fast moving and crowded world make aimless education intolerable."[2]

The present study, then, is conceived as a contribution to the task of developing a cross-national model that would enable the United States and five European democracies -- England, France, Italy, the Netherlands, and West Germany -- to make more adequate estimates of the effectiveness of their primary and secondary schools. The decision to adopt the cross-national approach was prompted by considerations of both "basic" and "applied" scholarship. The development of a body of comparative data on educational structures and functions is indispensable to the construction of systematic social theory on the relationship between education and society. From a more "practical" perspective, countries with some common and some divergent political and educational characteristics might derive mutual benefit from a close scrutiny of these patterns of resemblance and difference.

The choice of the particular six countries represented in this study was dictated in part by the desire to reduce the number of uncontrolled variables that might contaminate the results of the inquiry. In addition to comparable political structure and ideologies the countries share some important social and educational goals. Thus, for example, all depend on formal educational institutions as the main instruments of social continuity and peaceful change and seek to maximize cognitive abilities, develop proper attitudes toward citizenship, and prepare young people for vocational careers.

Yet, as the present inquiry clearly indicates, these educational systems differ, sometimes substantially, in the goals they seek and in the outcomes they apparently achieve. Thus it seemed that these nations exhibited the type of "limited diversity" which would make comparative analysis both feasible and fruitful.

As it turned out the international aspect of this inquiry proved somewhat disappointing. For a variety of reasons -- linguistic barriers, physical distance, differential research traditions, diverse intellectual styles -- the American and European research teams failed to achieve a full meeting of the minds. The materials furnished by our foreign colleagues while frequently illuminating in themselves could not be integrated in a single coherent orientation to our task. There were notable exceptions to this generalization -- the English contribution is most like ours in mood and substance -- but we make no claim to having made any appreciable advance in the field of comparative education.

From the outset it was clear that the development of a model capable of dealing with the educational systems of complex democratic societies would ideally involve extensive attention to three clusters of interrelated tasks.

1. The first set of problems consists of identifying a) those goals of a society and individuals that they seek to accomplish through the educational system; b) those means within the educational system that are designed to achieve the desired outcomes; c) the disparities that exist between avowed goals and actual achievements.

2. A second set of issues involves a) the perception of the disparities; b) the explanations that are offered for their existence; c) the actual reasons that account for them; and d) the sources of dissonance between proffered explanations and actual reasons.

3. A third battery of variables includes a) the strains and tensions that are generated by the disparities between goals and achievements; b) the mechanisms that emerge to reduce such strains including efforts to eliminate the disparities or to modify goals, means, or both; and c) the consequences for the next stages of educational goal-seeking.

As the body of this report should make clear it is not now possible to offer any authoritative solution to the first array of problems, much less to succeeding steps in the development of a comprehensive model of educational systems. We have succeeded mainly in specifying the necessary conditions for evaluating the effectiveness of formal schooling and in identifying the principal obstacles that obstruct the achievement of this aim. Although we experimented extensively with analytic schemes which are discussed in detail in a later chapter, the present volume should be viewed primarily as a work of meta-theory. As will become apparent, we did not construct a viable taxonomy and our work, to this degree, must be judged a failure. At the same time we are persuaded for reasons that we hope transcend vanity that the errors in the present undertaking are fruitful and that the problem of creating a useful educational model is now closer to solution.

Speaking very generally, the process of evaluation consists of 1) determining the correspondence between a goal and an outcome and 2) assessing the contribution of the educational system to the observed effect. The first three chapters of this report show that all of the key terms in this formulation currently create severe difficulties; the goals of education can rarely be clearly discerned; the outcomes are seldom properly specified; and it is hardly ever possible to demonstrate the independent influence exerted by clearly identifiable features of the educational system on man and events.

The primary function of the early chapters is to expose the flaws and inadequacies of the available literature, and to indicate the technical requirements which must be met in order to achieve a satisfactory model. Our principal conclusion, however, is that quite aside from the innumerable methodological difficulties of dealing with the full array of relevant variables and despite the obvious long-range advantages of a full-scale model, the first, and perhaps only, practical step that can now be undertaken is an inventory of incomes rather than a more complete scheme of evaluation which also emphasizes the role of educational goals. Stated quite simply in a democratic society numerous publics have a legitimate interest in the education of children -- educators, parents, children, corporations, unions, political leaders, and so on -- and it is highly improbable that anything approximating a meaningful consensus on goals can now be achieved. At the same time it is possible to persuade large numbers of people that systematic knowledge of what schools do or do not achieve is woefully inadequate. There can be no rational pressure -- whatever its motivation or rationale -- for changes in educational policy in the absence of any evidence of what happens to a child as he proceeds through his school career.

This consideration led us to attempt two major tasks: the development of 1) a taxonomy for the classification of claimed and confirmed outcomes and 2) a comprehensive measure analogous to the Gross National Product which has proved so useful to economists. We experienced only indifferent success in the pursuit of these objectives.

The taxonomy of outcomes consists of three major elements: 1) domains, major institutional roles and nonrole-structured activies, 2) aspects, subdivisions of domains, and 3) facets, modalities of knowing, valuing, and doing. The scheme permits classification of some 263 possible types of outcome. The advantage of this model is that it becomes a device for the location of any and all educational outcomes, and permits the perception of gaps in knowledge, comparisons with stated intentions, and other modes of analysis which are mentioned in the appropriate section of this work. Unfortunately, field trials reported in the appendices by Smith and Smith, by Krauss and Waldron, and by our foreign colleagues demonstrated the difficulty of making unambiguous entries.

James McGregor's insightful chapter shows that it is not possible to develop some such indicator as a gross educational product in the absence of uniform methods of valuation or "calculable terms of trade." Instead we are led to conclude that there are a variety of perspectives, which might serve as models for the organization of outcomes. These include models of 1) cost/efficiency, 2) resource development, 3) social (integration-disintegration and mobility-equality dimensions), 4) psychological models (personal and additive personal focus), and 5) humanistic models (religious, cultural, and citizenship focus).

The final chapter then ponders the lessons learned in the process of conducting this inquiry. We specify the components of a model for the measurement of the effectiveness of a school system whose major components are outcomes, reactions by diverse publics, average gain per student, community input and indices of social quality.

This brief map of our sojourn is in the manner of introductory chapters more orderly than it appeared while we were in the process of travel. This project involved the collaboration of a dozen foreign colleagues who served in the international phase of the project as supervisors, authors, or as conferees at a week-long session in which we compared findings; six American research associates or assistants who at various times wrote chapters, prepared appendices, acted as coders, experimented with provisional schemes, and compiled bibliographies; and of course the principal investigators who supervised the entire enterprise and made their contributions from generally similar perspectives which nevertheless did not preclude occasionally serious differences of opinion. The work as a whole represents their joint thought and planning. Marvin Bressler is primarily responsible for the introduction, Chapters 3 and 4 and parts of the final chapter; Melvin Tumin is primarily responsible for Chapters 2 , 5 , 6, and most of Chapter 8 ; and James McGregor is primarily responsible for Chapter 7 . The work of other persons appears in the appendices and their names appear with their contributions.

It would be difficult to impose anything approximating monolithic order on so many people, engaged in a complex division of labor, in a project which with planned interruptions took place over a span of two years and which addressed itself to novel and uncharted issues. It would in any case have been undesirable to insist on a "party line" in an area where so little is certain. The problem of assessing the effectiveness of educational systems is sufficiently amorphous to invite strong and divergent points of view.

It is not surprising, therefore, that this report has its share of inconsistencies and differences of emphasis. We have recorded in the appropriate chapters and reproduced in the appendices the demurrers of our European colleagues and American contributors. All of these pluralistic impulses were further encouraged by the fact that many of the same problems appear in modified form in several chapters which were revised because of consultation and a changed outlook while other authors retained their original points of view.

Nevertheless, there is a basic unity in our collective effort that transcends individual differences. We are persuaded that it is of the highest importance to devise better ways and means than now exist to measure how effectively educational systems achieve the tasks that are assigned to them in the United States and abroad. Moreover, although we are painfully aware that our triumphs have been few and our defeats have been many, we are content that the development of an analytic device that is capable of systematically ordering, recording, and storing theoretically adequate and operationally defined components of a comprehensive educational model will in due course prove to be within the province of human ingenuity.

References

1. A Plan for Evaluating the Quality of Educational Programs in Pennsylvania, Harrisburg, Pa.: State Board of Education, 1965.

2. Henry S. Dyer, "The Discovery and Development of Educational Goals, Summary," p. 1.

B. FINDINGS AND ANALYSIS

2. Education as a Social System

Functions, Structures, and Strains

It is elementary wisdom for any enterprise that risks valuable resources to secure certain desired ends to secure the most reliable possible estimates of whether the ends are being achieved, by the means employed, and the costs anticipated. In its basic terms, such an estimate is implied in the term evaluation.

Certain types of organization are advantaged in this regard by the fact that the criteria of "success" are built into the nature of the enterprise itself. Thus, in a gross and crude way, profit-oriented business enterprises know roughly whether they are succeeding or failing in their intentions by the margin of net profit they compile. This is admittedly a very crude measure, ignoring as it does, for instance, such problems as whether the profit sould have been or could have been much greater, given the costs incurred. But it serves at least as some measure of "outcome" which most other types of organized human effort cannot so easily call upon.

Consider for instance the problems faced by a church. How is one to determine whether the church in its totality of activities has been successful? What does successful mean? How could it be measured? How do you analyze the costs and inputs? How much of what type of success, with whom, should be expected in view of the input? If the parish is swollen with members during Sunday services, is this a signal of success? But suppose the same parishioners engage in daily activities directly counter to church preaching? How successful then can the church be said to be?

Consider, too, a peace-time army. What could determine whether it is being "successful"? What could it mean? If the test of the success of peace-time training is performance during war, but if the army loses a war to a greater power, does that mean the army is a failure? By what standards? Was it reasonable to expect the army to win? And if it never engages in warfare, how does one know if things are being done correctly? By what standards?

Consider, too, the problems of a baseball team that is a member of a ten-team league. If only the team that comes out on top is considered successful, then of course all nine other teams have been failures. The application of such an extreme criterion is likely to be disastrous to the morale and conduct of the members of the other nine teams. Obviously, other cutting points of success and failure have to be established; reasonable expectations have to be set up in view of inputs and probabilities.

No further examples are needed to illustrate the difficulties likely to be encountered by any organization as it seeks to evaluate its effectiveness. Admittedly, then, evaluation appears to be a most difficult task.

But any enterprise that pretends to rationality, that is forced to provide some kind of accounting of its operations, and that is beholden to one or another public for its successful conduct, requires regular and systematic evaluation.

Educational systems are no exception to this generalization. Whether public or private, but especially in the former case, public accountability is an integral part of the system. Moreover, substantial and valued resources are invested in educational systems; high expectations for important outcomes are held by the relevant publics; significant consequences for the "subjects," i. e. the students, are to be anticipated.

On the surface of the matter, "evaluation" seems to be a highly developed specialty in the field of education. Consider, in this regard, the presence of several national testing services; the availability of hundreds if not thousands of testing devices, standardized and other; the publicity given to national norms of achievement by which different school systems can measure and compare the outcomes of their own students on standardized test inventories; the proliferation of testing devices through every level from kindergarten to professional school. All these might be taken as reasonable signs that there is very great concern among the publics who make up the educational enterprise for sound indicators of how well their various schools are doing.

No one can gainsay that concern. It is to be found expressed in one way or another in the individual homes of the students; at PTA meetings; teacher caucuses; superintendents' conferences; statewide assemblies; and even at summit conferences at the White House. It is to be found concretely embodied in the Education Acts of 1964 with their specific and explicit concern for improvement of the quality of education and the equality with which this quality is distributed.

But one must address to all this expressed concern and to all the devices for testing outcomes that are currently being employed or advocated the question of whether the concepts are properly focused. Are they too narrow? too broad? tangential? Are the testing devices and instruments measuring what "should" be measured?

The introduction of the term "should" implies, of course, that there is a correct and fitting range of concern that should be embodied in evaluative procedures, and that their fitness and correctness can be shown by some persuasive logic. It stands to reason that one will seek to evaluate or measure those outcomes which he believes the enterprise ought to be achieving. It therefore follows that first consideration in the development of an evaluative scheme must be given to statements of objectives. As we shall soon see, for a variety of reasons this is much easier said than done.

Let us consider the problem first in its most general terms. How does one know what are the objectives of an enterprise? If there are a number of actors at different positions in the structure of the enterprise, and each has his own set of interests, expectations, and roles to perform, and hence each has his own specific and unique sets of possible gratifications, whose version of the goals of the enterprise should be followed?

From the point of view of the social scientist concerned with evaluation, it is impossible to take sides in the dispute as to which public's version shall be attended to, or which index shall be used; unless there is some larger conception of the educational enterprise to which the criteria of decision and judgment can be related.

Admittedly there will be disagreement among various sociologists as to what "education," as a set of social actions, comprises. Differing definitions of education, and the attendant variant specifications of the boundaries of that social institution, will thus lead to different sets of constituent objectives, whose achievements are to be measured in any evaluation of the effectiveness of a school system. But some definition must be present at the outset to provide a rationale for all that follows.

One can even state some of the criteria that such a definition ought to satisfy. At a minimum, the definition ought to make it possible to "code" any activity, reasonably unambiguously, with regard to whether it falls within the boundaries established by the definition.

Unhappily the definitions of social institutions rarely make this kind of pointed coding possible. For even in analytic terms we are often required, for example, to say nothing more specific than that the economy is that aspect of a social system which is primarily concerned with the production and distribution of goods and services. When then we turn to lists of various activities to decide whether they are properly called economic or not, we find that they are enmeshed with numerous other institutional activities, such as political, familial, and educational, and our decision as to whether the activity is properly called economic is thus forced to rely on a fair amount of subjective judgment. Thus, the investment of school bonds, for purposes of producing school-related revenues, can be seen either as an economic aspect of educational activity or an educationally oriented aspect of economic activity.

In any event, the establishment of "coder reliability" is a matter of internal consistency and has nothing to do with external validity. That is, even if numerous judges independently coded materials in the identical way, it could still be problematic as to whether what they were coding so well was "really" what one wanted information about. There is a common tendency at this point to utter an "ejaculation of despair" to the effect that "everyone knows what education is anyway." But, if that were so, there would be no debates about who, why, what, when, where, and how in education today. Such debates and disagreements sometimes seem more frequent and basic than the agreements.

Clearly, then, a definition of education as an enterprise is required. This definition must coincide to some effective degree with common perceptions of what education is about, or else it will serve little purpose. Additionally, the definition has to specify the boundaries of the educational enterprise clearly enough to indicate the points at which external forces are penetrating the system, and conversely, the lines at which educational forces are penetrating into or intersecting non-educational systems.

Other specifications will be cited later as they become appropriate for the level of specificity, or the particular kind of educational system, that may come into consideration. For now it will suffice to state the more general terms of such a definition, aware as we are that, while all definitions are ultimately arbitrary, they must be useful.

A. The Boundaries of the Educational System

The term "system" gives an initial set of guidelines to the definition of education for it requires that the definition be put in "system" terms and that it accept certain limitations imposed upon it by virtue of the fact that it must always be seen as a part of a larger system, called the social system, within which it operates.

The human social system consists broadly speaking of all these structures and activities relevant to the maintenance of organized, i. e. patterned, human life over a period of more than one generation. The relevant structures, among others, are those concerned with production and distribution of goods and services, the maintenance of order, the allocation of members to their status or slots and the inculcation of their roles.

From the point of view of the human social system, one of the crucial and universal functions of all such systems is that of transmitting from one generation to another the skills, ideas, values, knowledge, and beliefs considered by the older generation to be necessary for the survival and functioning of the new generation. The necessity for such transmission arises out of the fact that the biological inheritance of the human being is simply insufficient to permit any human to survive and function without learning how to do so. The second imperative arises from the fact that if the individual is to live with others, he has to learn the rules and the sanctions of that pattern of coexistence, and he has to find an appropriate place within the division of labor.

Two things, then, have to be taught to the newborn infant: how to survive and how to survive in a patterned and acceptable way -- acceptable, that is, to the others with whom he must relate if he is to survive.

In an enthusiasm for widening the domain of "education," some scholars have argued that education is nothing more nor less than the totality of all organized social life, since in fact learning is always going on, all human experiences are "instructive," or can be, and every interchange between humans is potentially educational in that it is either concerned with or consequential for the transmission or acquisition of knowledge, skills, etc. By this token, however, all experiences are also always economic, always political, etc. For in some sense, direct or indirect, all human interchanges are consequential to some degree, for all human activities of any and all kinds. It is obviously beyond reason or utility, then, to seek the widest possible extension of the definition of education.

We deliberately confine it, therefore, to those structures and activities that are primarily concerned with or focused upon the transmission of knowledge and skills. We impose one further limit, though fluidly rather than rigidly so. It is to the effect that we are concerned with the statistically most frequent situation of a legally "underage" person being taught by a person legally "of age." Underage here refers to that situation in life where the individual is adjudged young enough to require further instruction in necessary values and skills on the one hand, and young enough to legitimate his being subjugated in the process to the will of the older person involved.

There are, of course, exceptions to this age distinction. It does not always happen that the teacher is older than the pupil. But in such cases, either the teacher is "functionally" older, in the special sense of possessing skills and knowledge that the other person does not possess, or the situation is not primarily educational, as we have so far defined that term. Thus, a younger woman instructing an older man in certain refinements in dallying would surely be "teaching" the man, but this "teaching" would not fall within the range of the knowledge and values and skills that are today believed to be required for acceptable adult functioning, however enlightening and pleasurable they might be.

Several basic elements now are available to guide us in the definition of education: 1) education always occurs in a context of a larger social system; 2) the continuity of this system over time requires that new members of the society brought in by birth or reproduction be taught basic requirements of patterned survival; 3) some legitimate agent for such teaching is always designated by the social system.

We can join these and other elements in the following statements:

1. Education always involves transaction between a person, usually an adult, defined and designated officially as a teacher, and a person, usually a non-adult, defined as a student. The adult - non-adult distinction need not be a formal age distinction, but there must be an acknowledged difference in the knowledge about the relevant subject or experience that are to be communicated.

2. The main purpose of the transaction is the transmission by the teacher, or with his facilitation, of a set of understandings, with varying affective, conative, and cognitive elements, which a) are considered officially important by a legitimate authority for the child or learner to know, and b) about which there is presumably a defined content which can be transmitted. Legitimate and valued content are the characteristic substance of the educational transaction.

3. The central purpose of education is the rendering fit for adult life of the unfit child. The purpose of this transaction is seen as conducing to the benefit both of the learner himself and the society or community in which the transaction is taking place.

(Different communities vary with regard to their evaluations of the final locus of the benefit of the educational transaction. Some emphasize the ultimate benefit to society more than to the learner. But everywhere some division of benefit is allocated among both learner and the society.)

4. The legitimacy of community support of this transaction is defined in terms of the likelihood that the learner will become an adult functioning member of the community. Insofar, that is, as he is the "learner" in the educational transaction, he is defined by the community as a "dependent" and as less than a full functioning member of the community. His "learning" is thus seen as one of the things needed for him to validate his claim to full membership.

5. Insofar as the teaching agent is seen as the representative of the community, and as the individual with the specialized content that is to be transmitted, he is always assumed to be endowed with legitimate, and even functionally demonstrable powers and authority to prescribe the mode by which the transaction shall take place. Either the teacher-agent is delegated these powers or he is designated as the person who fulfills the "plan" of education for the community under its guidance and direction.

6. The right to designate who shall conduct the transaction, under what circumstances, and by what means, is justified as belonging to the community by virtue a) of its interest in the outcomes and b) its provision of the resources for the transaction.

7. The community, acting as a body or through a designated agency, prescribes not only the actors and the content of the transaction, but also the amount of time that shall be devoted to the process; the physical place in which the transaction shall occur; the internal rhythms of time and personnel; the kinds of supporting materials and experiences that shall be designated as officially necessary and relevant. Communities vary in whom they designate as relevant to the decision-making process regarding these dimensions of the transaction. But the right to so designate is presumed to lie in the resource-providing community within which the system is located. Norms of time, place, and situation regarding education are found in all societies, even those without formal systems of education.

8. The community always prescribes, directly or indirectly, a body or set of moral principles of conduct that are meant both to govern the transaction as it takes place, and that are to be imparted to the learner by the teacher, either as central contents of the transaction (along with other contents) or as incidental and instrumental to the learning of other content. These rules govern behavior immediately relevant to the educational transaction and more ultimately relevant to the conduct of the learner, then, and later in the community at large.

9. In his role as agent of the community, the teacher is not seen as having any needs which must be served by the transaction except insofar as they can be shown to be relevant to the outcome of the transaction, and as specifiable in terms of gains to the learner or to the community or both. The teacher, in short, is an agent with a specified role. Tolerance for deviation from this role because of differences in personality or other such factors is always at the discretion of the funding and legitimizing community. The range of such permitted deviations is variable from one community to the next, but is always justified in terms of toler-

able deviance, or valuable individuality. It is never justified in terms of the needs of the teacher as a person.

10. All of the content involved in the transaction is seen as belonging to two sets of interlocking structures. On the one hand, each of the separate "subjects" being taught is seen as having a structure which can be broken into steps, from lesser to greater difficulty and complexity, and progression along these steps to the completion of the skill or understanding is seen as measurable. In the second instance, all the subjects taught are seen as constituting a total structure of things important to learn because they are collectively valuable to the learner, or to the society, or to both, even though their connections with each other may not be demonstrable. In short, there is always a vision of an "educated man" that guides the community in the selection of the content of the educational transaction, and that enables it to set minimum levels that any acceptable individual must presumably reach.

11. All educational transactions are seen as capable of being so defined that norms of achievement at specified normative intervals can be set by the community. Expectations by age or number of years of exposure or some other such indicator are always present.

12. As in all systems, there are specifiable limits of tolerance for deviation from expectations by both the learners and teachers. To counter possible deviations beforehand, or restrain them when they occur, there are always available certain specified sanctions that may be invoked by the legitimate authorities. All educational systems allocate to the teacher himself certain sanctioning prerogatives. More complex systems may require the teacher to refer questions of sanction to more remote authorities. But sanctions are always present; they are known, patterned, legitimate; and they are expected to be confined to the task of reducing or eliminating the obstacles to the envisioned or desirable process of transmission. In that regard, they are "task-specific." They are more narrowly constrained, in fact, than the sanctions permitted or endorsed for deviations from role expectations in other systems such as the political or the economic. The restraint upon these sanctions -- their confinement to "task-specific outcomes -- is because it is felt that the learner, the teacher, and the social system in which the learning-teaching is going on will lose something valuable if the deviations continue and if learning-teaching is halted. The object of the sanctions is to get the educational process back to normal course as quickly and as effectively as possible. In this regard, sanctions tend to be much "milder" and narrower than, for instance, criminal sanctions for violation of property codes. The demand for the most immediate possible restoration of equilibrium in the educational process determines these particularities.

So far we have specified some of the following things about education as a system:

1. the minimum cast of actors involved;
2. the essential relationships between them;
3. the common purposes ascribed to the relationship;
4. the "unit" of transaction (educational content vs. money vs. power vs. privilege, etc.)
5. who has the power to determine how the relationship shall be conducted;
6. the justifications for this distribution of power;
7. who is responsible for providing the resources for the transaction and how is this responsibility justified;
8. the limits of tolerable deviation, who has the sanctioning power, and what are the constraints on those sanctions.

B. The Uniqueness of the Educational System

The combination of these twelve characteristics makes the educational transaction unique by comparison with any other we could specify. Educational relationships, in the framework of these rules of transaction, are specifiably different from political relationships, or sexual, or familial, though in each of these some education may take place, and sometimes one or another member of the family may be designated by the community as the official agent or teacher. The teacher-learner relationship may thus be one role-facet of the parent-child relationship. But even when this is the case, the relationship between parent and child when they are functioning as teacher and pupil is different from their relationship when they are out on a hunt or fishing expedition together (though there may be educational aspects to each of these ventures as well); or when they are eating at their common household table; or when they engage in religious devotions together.

The difference between education and these other activities is in the degree of focus upon a central intention. The educational focus is upon the intention of transmitting valued knowledge and skills from one generation to another, to prepare the younger generation to be able to function as adults, and thus to benefit the larger society by having such trained functionaries.

In specifying the essential characteristic of the educational enterprise, no specific structures as such were mentioned. That is, we have not spoken of the educational system in France, or Tobago, or Malaya. Rather, we have kept in mind the entire range of possible structures within when these educational goals are sought and the correlative fitting actions taken toward these goals by people in the roles of teacher and learner.

In speaking of the "community" and its controls and delegated or exercised powers, and its provision of resources and setting of norms, we may have unwittingly suggested a highly structured situation. But our intention was to include in this description the educational transaction of the Eskimo father teaching his son how to spear fish as well as the highly sophisticated urban-dwelling teacher discussing the "Great Books of Western Civilization" with his college-bound students. Without unduly straining the terms described, both these situations and all those in between these extremes are meant to be included. In the Eskimo case, the "community" is an understood community, the one to which the father and son both are committed by identity and bonds of loyalty, and from whom they, like other Eskimoes, each at his own fishing hole, far separated in space, derive for themselves norms and judgments. The formality or informality of the contact within the community, and of the lines of connection with its norms and sanctions do not matter at this stage in the development of the boundaries and content of the educational system. These specifications become important only later when one is attempting to distinguish various kinds of concrete educational systems.

C. Toward a More Complex Model of the Educational System

Thus far we have talked of an educational system in terms of the simplest possible model -- that of a teacher-learner relationship, with only the two actors involved. Most systems of the world would not correspond to this model. Rather, most are somewhat more complex.

Now then we may add the complexities which are introduced only when we increase the numbers of actors playing the role of teachers and learners.

1. Complexities caused by added numbers of participants

If we keep in mind the model of educational systems in modern, industrial societies, the following elements seem to be characteristic of all such systems:

1. There are now diverse and numerous actors called "teachers" who are needed to staff the system.

2. Specialized training must be provided so that they can be adequate to the demands upon them.

3. That training has to be relatively coordinated -- within the domains of the same society -- so as to ensure a working consensus on the purposes of the educational transaction, the content to be transmitted, the methods to be used, the time schedule to be kept, the evaluation techniques to be employed, and the competence of the teachers themselves.

4. That system of training, and its predictable and measurable outcomes, has to be coordinated with the perception of the purposes and content of education held by the various communities in the society who are going to employ the teachers.

5. Formal arrangements must be made regarding the governance of the educational operations. Firm lines of allocation of power to make decisions must be established. Where bureaucratic hierarchies are needed, their rationale and structures must be agreed upon.

The requirements just specified indicate that many new kinds of actors or role players have been brought into the educational enterprise. It has become complex in its needs and operations and hence a division of labor has had to be set up. These new status or slots in the division of labor involve recruiting people for non-educational functions required to serve the educational enterprise, including suppliers of goods and services ranging from coal or oil for the furnaces of buildings, to paper for the classes, and food for the cafeterias, and lawyers for the flotation of bonds or loans.

This is one level of complexity -- the increase in the number of people to be served in strictly educational ways -- where education is defined as we have done earlier.

2. Complexities caused by new relationships

As a result of these new needs for training and governance, whole new sets of relationships emerge which become integral parts of the educational system and which can come into conflict with each other unless otherwise managed. These include the following:

1. between apprentice teachers and teachers of teachers;

2. between formal government authorities and the teachers of teachers at the specialized institutes;

3. between the formal governance authorities and the parents of the children who are to be taught by the teachers;

4. between the teachers and their superiors in the hierarchies or bureaucracies set up to manage educational system affairs, e. g. members of the ministry of education or the school boards, or whatever;

5. between the parents and the teachers who now play the role parents used to play in the educational system;

6. between the students and the teachers who now are no longer members of their families or in multi-role connection with them;

7. among the various teachers within the framework of their own teaching institutions -- when more than one specialty is taught by more than one specialist;

8. between the formal governance authorities and the manufacturers and suppliers of goods and services required to provide the schools with the necessary resources;

9. between the teachers of teachers and that new band of specialists who innovate materials, methods, and processes regarding education, i. e. the research specialists in the field of education.

There are, of course, other role players and other sets of relationships which develop when one adds more students and teachers. But those just cited will suffice to indicate the kinds of complexities that are introduced simply when relationships are increased.

3. Complexities caused by inter-system affiliations

Another layer of complexity is introduced into education when any system "moves" out of a condition of total local sovereignty and becomes part of a network of educational institutions, arranged in a hierarchy of coordinated authority. Necessarily there is involved some surrender of sovereignty and autonomy by each local unit, and the development of new systems of relationship to manage the interaction between the local and the federation of which it is a part. The connections between local, county, state, and federal educational networks will serve as an appropriate model here.

In this new complexity it is not that new purposes have been added to the core educational purposes. Rather larger numbers of people in various communities have, for one reason or another, pooled their purposes and their resources, and have thereby set up certain requirements of coordination and governance that were not present earlier. This is not simply an enlargement of the prior increase in numbers. Under that simpler condition no structural realignments of authority and autonomy were required. Only coordination among diverse agents of the same sovereign power serving the same simple purposes had to be realized.

4. Complexities caused by addition of functions

Our model of education has grown thus far to include large numbers of actors, organized together into a network of reciprocally supporting efforts, operating out of various centers with varying degrees of autonomy and sovereignty, but still focusing on certain specialized and limited purposes.

In the simplest educational systems, the purposes of education seem quite clear and unquestionable to the actors involved. There are no debates in "Pango-Pango" as to what should the father teach his children. Above all, there are no debates as to whether morals and values ought to be taught, along with knowledge and skills, and no questions regarding the utility of what is to be taught -- not until alternative ways of possible lives for the children become real options to them or to the culture at large.

But things do not remain that simple and "uncontaminated" for very long, at least not very long after the society in which the education is transpiring comes into effective contact with the impulses of modernization and industrialization.

Under the impact of this kind of contact, several basic changes in societies occur which are directly relevant for the educational system:

1. The division of labor becomes such that training for specialized and unequally-valued occupations becomes a necessary part of the schools' efforts.

2. The family surrenders most of its educational functions and assigns them to the schools. This happens partly because the parents are themselves too involved in the making of a living to be able to perform the educational functions, and partly because the content of education at certain levels becomes too complex for the parents to be adequate.

3. Diverse careers become possible for the students so that new specialties are developed in the school requiring new specialists.

4. The diversity of talent among students for certain specifiable tasks in the adult world now becomes a salient consideration and patterns of selective recruitment, training, and placement of students become primary considerations for the educational system.

5. The schools are required to take on a diversity of new functions not simply in the traditional pedagogic sense but because of the loss or surrender by the family of its capabilities or interests in these concerns. The schools become lunchrooms, recreation halls, laces of religious devotion and instruction, mental health screening centers, and vocational guidance counseling offices, among other things. These specialties fall upon the schools as the need for them emerges, mostly because they arise as needs of the school-age youth. The fact that such youth are "captive" in school centers for specified hours a day makes the schools the most "natural" target. The alternative would be to allocate each of these non-pedagogic functions to specialized agencies outside the school, and set up a network of coordination between them and the schools, insofar as it seemed necessary and in the best interests of their common subjects, the school-age children. In fact, there are numerous arrangements, including in-school and out-of-school agencies, and varying kinds of coop-eration. Another alternative is for the schools to refuse to perform the obviously needed functions. But when the need is recognized, and when there is some action toward meeting the need, the school more often than not becomes the agency to whom the servicing of the new need tends to be assigned.

It is important to note here that the functions taken on by the schools under the pressure of modernizing and industrializing influences can all be fitted into the original simple framework of "knowledge, skills, values, ideas, and beliefs." But this is a semantic concealment of genuine change. Among other things the unavoidable augmentation of knowledge and skills results in the unavoidable loss of competence of the parents or teachers to acquire and impart these. Additionally, the new dimensions of effort that emerge, e.g. vocational counseling and mental health activities, are "verbally" inherent in the mandate of the school to "prepare the child for adult life," but were never visibly required before the specialized division of labor characteristic of modern industrial society became a reality.

Some of the additional complexities introduced into the system by the addition of new functions can be summarized as follows:

1. new roles for new specialists;
2. new relationships between the old-line specialists and the new;
3. new relationships between the new specialists and the homes and families of the children, and the children themselves;
4. new sources of funding and financing required;
5. alterations in the structure of competence and authority in the traditional hierarchy developed around the more limited sets of purposes;
6. new relationships with funding and sponsoring authorities outside the school who have now been brought into the educational network;
7. new interrelationships with newly relevant professional organizations, such as medicine, engineering, law, and accounting;
8. new relationships with organizations such as trade unions.

5. Complexities caused by education of children within a democratic school system for adult function in a democratic society

No attention has yet been paid in the development of a model of an educational system to those complexities which arise when a) the system operates in accordance with democratic canons of decision-making and b) when it seeks to prepare children for life in a democratic society. These are very specific kinds of cultural themes that can give a very special shape to the educational system, especially if the system takes seriously the implications of being democratic itself, on the one hand, and of training its children for democracy, on the other. The complexities that arise under those circumstances are over and above all those already specified. The difference can be seen best when one imaginatively compares the problems facing such a system with those confronted by the rulers of a system which is unabashedly autocratic in its own operations and which, with equal devotion, aims at preparing its children for life in an autocratic or totalitarian society.

It is important at the outset to distinguish the two "democratic" themes of which speak here -- the democratic governance of the system itself and the training of the children for democracy. In theory, the combination of these two themes yield four possibilities, as shown in the following diagram:

		Type of System Decision-Making	
		Democratic	Non-Democratic
Type of Educational Training of Children	Democratic	a	c
	Non-Democratic	b	d

No empirical society is a total and perfect fulfillment of any of the four types indicated above, but some systems come closer than others. The American system in general comes closest of any in the known world to type "a." Selected communities in the United States come close to type "b" insofar as they may go through all the mechanisms and processes of democracy to decide to teach their children to respond to themes of hierarchy, inequality, discipline, authority, etc. Type "c" represents the kind of educational system one would expect, under the best possible circumstances, from societies "in transition," where the transitional period of autocratic governance is seen as the necessary and unavoidable condition for survival but where the aim, nominally at least, is to create a new kind of democratic society. Type "d" is the classic case of education in a totalitarian society where the clearcut intent is to educate for totalitarian conduct.

Type "a," the system marked by democratically-decided democratic content, is probably the most difficult to create and operate. Democratic decision-making is the most laborious and time-consuming of all modes of decision-making, and training children for democracy involves modeling that democratic pattern in the actual conduct of the schools, and entertaining the highest levels of openess for contingency, disagreement, and alternative theories and viewpoints.

When a system decides to be democratic in its structure and contents it opts, willy-nilly, to break down the traditional authority of the teacher and the authority of the traditional subject matter. The body of "received knowledge" is no longer sufficient; the modes of selection of appropriate materials are no longer as relevant or justifiable. The alternative courses in life are no longer as restricted or constrained. The diversities of talents among students are no longer as easily arranged and utilized in hierarchical fashion; inequalities in educational outcomes are no longer as easily rationalizable. The relationships among student, parent, and teacher are no longer capable of articulation in the same way. These and numerous other changes are required when one changes not only to a complex system fit for modern industrial life but one also fit for life in a democratic society.

Perhaps just as important is the fact that almost always when systems move toward democratic structure and content they also embrace the notion of the naturalness and need for change, and continuous change at that. The inherent instability of social life, and the positive desirability of constant change to meet newly created needs and desires become central orientations of such a system. At that point the character of education becomes radically changed. For now it is no longer the continuity of valued traditions that serves as the center of school content. Rather it is the continuous revision of traditional values and valued traditions to make them apposite new situations that becomes the keynote. The great _sturm und drang_ now being experienced by even the most enlightened schools in the United States as they try to define, comprehend, and begin to implement to notion of "education for change" is evident testimony to the difficulties created when the old verities are seen as no longer sufficient. Moreover, those very few pioneer schools who have only recently decided to see what it might be like both to provide models of democratic life on their campuses and to allow democracy to be the rule of conduct in school affairs are experiencing what might more properly be called birth pangs rather than adolescent _sturm und drang_.

It is clear, therefore, that democratically decided educational content for democratic life is the most difficult of conceivable school systems to manage effectively.

Summary

An effort has been made to set certain analytical boundaries upon the system of human effort called "education." This has involved specifying the key personnel, the relationships among them as they move toward agreed upon goals, the nature of those goals, the lines of allocation of responsibility and of provision of resources, the power relationships that arise, the permitted deviation, the mechanisms and agencies for sanctions against deviation.

Once these core elements, characteristic of all systems of education, were specified, the model of the educational system so described was then elaborated to take account of the new features and system complexities that arise when 1) numbers of actors are multiplied; 2) new relationships must be structured and maintained within the system; 3) new relationships with bordering systems are developed; 4) new functions are added to the charter of the educational system; and 5) when the system attempts to operate in accordance with democratic canons of procedure and to instill the knowledge and capacity requisite for democratic living in its students.

With these boundaries, structures, and themes now stated, it is relevant to consider the strain and stresses that are encountered by educational systems, at the various specified levels of complexity.

Part II. Strains in Educational Systems

Every form of collective life is inherently stressful. Material resources are often inadequate; the knowledge required for rational functioning is often unavoidable; the needed cooperation among diverse personnel is often lacking because the cross-purposes of the various actors are too divergent to permit easy reconciliation.

Education is no less stressful than other collective efforts of man, and in some special ways may be more stressful. At the very least, it exhibits a number of strains that seem unique to it, given its essential purposes and its unavoidable minimal structure.

In outlying these strains, it will be useful to follow the same structure that was used in developing the model -- first, a core system and then take account of complexities that arise because of increased numbers, diverse centers of authority, increased range of functions, and the focus upon democratic governance and content.

We consider first a sample of those kinds of strains that arise from the fact that the system is "educational." Here we take account only of the generic strains that can be found in any and all educational systems, regardless of simplicity or complexity.

A. Strains Common to All Systems of Education

1. Ambiguities and variations in content

We refer here to the fact that even in the simplest societies, where education consists of skills, values, attitudes, and beliefs that father transmits to son, and that elders transmit to the age-graded young men, the skills themselves, e. g. hunting or fishing or taking scalps, are subject to being learned in a variety of ways. There is no manual of specifications for these skills. Ways that seem suitable to the teacher may not be apt for the pupil. Or, the pupil may discover modes of performing the required skills which seem superior to those of the teacher, or at least more fit for the student. The "authority" of the teacher, based in large part on his being perceived as a repository and monopoly-holder of important skills, may thereby be undermined.

2. The pupil outdistances the teacher

In the same vein, the authority of the teacher may be undermined, especially when strength and vigor are requisite to the skill, by the rapidity with which the learner surpasses the teacher in the performance of the skill. Given the age and generational differences in teacher-pupil relationships, and given the crucial role of muscular strength, agility and endurance in non-industrial societies, this strain is ubiquitous.

3. Teacher A is better than teacher B

Another possible subversion of the authority of the teacher arises in the situation where unavoidably the learners compare themselves against each other in terms of the skills of their adult teachers. Variability in the adult population guarantees some invidious comparisons and the consequent lessening in the authoritative prestige of the less competent teachers. The industrial society equivalent of this occurs both in the schools, when comparisons are made among teachers, and in the peer-groups, when comparisons are made among fathers.

4. Insufficient resources

Because of lack of time, or energy, or interest, or knowledge or skill, or even of minimal material resources (a good bow or a good fishing net or a good horse), the education of the learner may be less adequate than is considered desirable. Especially in small agricultural or peasant or hunting communities, where such matters are visible, there may be unpleasant stressful repercussions on teacher and learner alike.

5. Generation gaps in value orientations

Under even the simplest social circumstances, differences in age and generation between teacher and learner may prove to be sources of strain in the educational transaction. The "internal" requirements and norms of the young peer group, with definitions of a desirable rhythm of work and play that differ from those of the adult generation cause difficulties everywhere.

6. Current responsibilities and deferred gratifications

The educational situation is one which everywhere, no matter what the kind of society, involves the young person in assuming certain "painful" responsibilities as a learner in return for which the reward, at least in part, is a gratification to be received later in the game. The learning of how to defer gratification is a crucial element in all effective education, even though in some social situations the gratification is deferred only momentarily and only in part. The inherent gratifications received "in process" or "along the way" may serve as compensating and motivating features. But there is always some element of "cost" along the way that must be traded off against future gratifications. Typically, the members of the adult generation, who have been through the process, are likely to give much more emphasis to later and deferred gratifications and hence to stress the worthwhileness of deferral of gratifications to an extent that cannot be shared by the younger learners.

7. The misfit of the personalities of teacher and learner

In any role-structured situation such as that of teacher-learner, none of the personality dimensions of either of the two actors, in principle, should enter relevantly into the transaction between them. But in all human interactions such "role-irrelevant" personality dimensions do in fact intrude. The amount of intrusion will increase in proportion to the number of other relationships in which those same actors are involved. It will also be greater where there are fewer formal specifications of the limits of the role. In the simpler society, then, where the adult plays the roles of father and teacher, breadwinner, religious senior, and many other things to his son, and where the son plays all the role counterparts of these various role-structures to his father, the intrusion of their "role-irrelevant" personalities is likely to be maximized. The compensating feature here may be the extent to which father and son are able to work out patterns of productive coexistence precisely because of the range, continuity, and scope of their interactions, so that there is a surplus beneficience that carries over from one transaction to another and may serve to compensate for inadequacies or strains that arise. But, of course, there is always the other possibility that each role-relationship will be progressively worsened, or lessened in its effectiveness, because of the spill-over of unsatisfactory relationships in other roles they share.

8. The intrusion of affect and other non-rational considerations

In any interaction between humans, however sharply defined and circumscribed, there is likely to be an important affective component. This is perhaps especially true where, from the viewpoint of the learner at least, the content of the transaction is crucial to him, his identity and ego-strength, and/or when he is subject to the consequential evaluation by the teacher, and when he normally has no recourse or appeal beyond that of the teacher's willingness to consider. In the father-son education relationship, the unavoidable affect inherent in the father-son aspect necessarily intrudes into their teacher-learner transaction. It may be supportive or destructive, but it is always problematic.

These are eight of many possible strains that can be and are encountered in educational situations, given only a minimal structure of two actors, in a situation devoid of formal codification or evaluation, and in which there is no formal network connecting such two-actor systems with others.

B. Strains Arising in Complex Democratic Systems

1. Strains caused by democratic governance and content

Most of the strains could properly be classified as disagreements among various actors in different roles (parents, teachers, students) regarding all those aspects of education which are subject to dispute: its purposes, its content, its methods, its best method of evaluation, its necessary resources, and the like.

Perhaps the most persisting and fundamental of these disagreements which seems to inhere in the very character of the system has to do with the content of education: what it is in terms of knowledge, values, beliefs, and skills that ought to be taught, at what level of competence, within what period of time, and for what later or immediate utility.

The emergence of widespread disagreement on the content of education, i.e. what should be taught, is because of three main sources, among others: a) the amorphousness and ambiguity regarding the purposes and goals of education, and hence regarding the content that might fulfill those purposes, especially of course in systems that have decided to train their children for democratic social life; b) the diversity and mixture of pedagogical and non-pedagogical, or curricular and non-curricular, goals of education that have come to characterize educational systems that profess to aim at fitting their students into a complex and rapidly changing world; and c) the diversity of importance attributed to various outcomes by the various publics connected with the running of democratic school systems, each of whom brings to its educational role-interaction a view of the educational process that is most consonant with its own place in the educational system (parents, taxpayers, professional teachers, professional administrators, etc.).

It is a commonplace that systems acquire special features of structure and function from their avowed purposes. Thus, a corporation engaged primarily in the business of manufacturing pipe fittings will be different in important regards from a corporation engaged primarily in recruiting entertainment talent for bookings in various theaters throughout the country. So, too, if the pipe-fitting corporation decides to take on the manufacture of newsprint, it will have to restructure itself in a number of ways to accomodate its organization to both purposes. Even if the manufacturing aspects of the enterprises are kept separate, a new structure of corporate reporting and authority must be created to bring the diverse enterprises under a common governance. Thus, when diversity and multiplicity of purpose are introduced into a previous one-purpose structure, the shape of the system is likely to alter in several significant regards.

If, however, it is characteristic of the system -- corporation or other -- that its purposes are stated ambiguously so that there are no specifiable provisions for the measurement of its achievement, it builds a high potential for disagreement regarding its effectiveness into the very nature of its operations.

Educational systems in democratic society are precisely of that kind. By announcing that their purpose is to prepare young men and women for "meaningful" adult functioning in a democratic society, they invite the maximum possible argument regarding their structure, their activities, and their achievement.

This is so simply because by comparison with any other stated "intention" there is more deliberate uncertainty inherent in the "democratic" goal than in any other. And this is so because, by the democratic norm of "process," the content of democratic education must always be in flux. Moreover, that flux tends to the maximum when the system encourages maximum participation by all relevant publics in the discussions and decisions as to what such a system ought to be doing to further itself and to insure its continuity over generational time. Uncertainty regarding content and methods are thus inherent in educational systems in democratic societies. Perhaps the most certain guideline for such educational systems is the requirement of training teachers and students to function effectively in situations of uncertainty and ambiguity and to benefit from the exchange of disagreements in the temporary resolutions of these disagreements.

Some will argue with the connotations of the term uncertainty and would prefer the term "provisionality." That might indeed be closer to the mark. For the essence of democratic educational systems is that none of the avowed purposes are viewed as anything but the most general guidelines. Correlatively, all instrumental processes or paths toward these ends that are prescribed at any time are seen, when the system is functioning adequately, as provisional and temporary, and subject to change when innovative research suggests better alternatives. In the interim, there is deliberate encouragement of such innovative research. In short, democratic education is constantly seeking ways in which to shake up, alter, and modify its existing procedures. In those regards, then, democratic education is inherently unstable with regard to its content, methods, and relationships.

The deliberate cultivation of disagreement through the explicit invitation to numbers of different publics to participate in decisions regarding the conduct of the schools is tantamount to the acceptance of a conflict model of social organization. This potential for conflict -- and especially between student and teacher -- is cultivated even further by the deliberate emphasis in modern democratic educational systems on the preparation of students for rapid social change and for effective participation in a world that will experience such change. Because of the generational difference between student and teacher, the emphasis on preparation for change, if implemented adequately, must result in a diminished quotient of authority for the teacher and in more active participation in decision-making by students. If these relationships are not changing in the ways just specified, it can be inferred that training the students for change is not proceeding as effectively as it might in the schools.

2. Strains arising from multiplicity of goals in democratic systems

A second set of characteristics of complex, democratic educational systems arises from the multiplicity of purposes which the mandate implied in the name of the system seems to suggest.

We refer here to the array of commonly affirmed goals of education in our own and comparable systems that go far beyond the traditional development of skills in basic literacy. These include a) training for such things as being able to live productively and peacefully with people who differ in certain basic characteristics, abilities, and opinions;

b) being motivated to participate effectively in the democratic process of decision-making through sustained activity in the decision-making process; and c) being able to acquire through participation in the educational system a sense of self-worth equal to that of all other participants.

No other educational system in the world has embraced these kinds of goals as fully as the American educational system. These goals present to the schools that accept them as legitimate and important certain most difficult problems connected with personnel, methods, content, resources, priorities of time and place, and methods of evaluation. Such schools become multi-purpose systems the achievement of whose ends is intensely difficult to estimate. The difficulties arise from the character of the goals themselves and from the crucial fact that the extent to which they have been achieved cannot be known with any firmness or certainty until the child is in effect an adult. Even then the accomplishment is most difficult to estimate and, it is even more difficult to assess the extent to which the achievement is due directly to the school (and which of its many facets), or was made in spite of the school, or if the school experience was relatively indifferent and inconsequential.

It is not multiplicity of goals alone that generates these difficulties. It is specifically the kinds of multiple goals that the democratic systems seem to have taken on for themselves. Multiplicity would present enough problems even if the goals were stated with crystal clarity and the operations for their measurement were prescribed clearly and beyond cavil. When these burdens are added to those previously described as arising from the democratic governance, with its unavoidable quotient of dissensus and conflict, it can be seen that democratic education is indeed a burdensome charge.

3. Strains arising from the presence of numerous, relevant actors in different places in the school system structure

It is axiomatic in the theory of systems that as the number of relevant actors or roles increases, the complexity of management increases, and the potential for conflict increases accordingly. Every actor in every system views the system from the special purchase or point of view that derives a) from his own position in the system and b) his own special ties to other systems outside. The student with his family ties; the teacher with his personal and professional connections; the principal with his special connections to the superintendent, the parents, and the school board; the parents with their special relationships in the nexus of the parent community: all these are actors who must disagree, at least in part, as to what the schools ought to be doing and how well they are doing it.

For example, the teacher who out of "professional pride" is "tough" in his evaluation of students must by that very token be a source of problem to the students on whom he has been tough. The students must in turn see the teacher and the educational process as working against their best interests and, even, against the best concept of effective education. The parents in turn must differ with their children regarding the adequacy of the schooling being forced in the nature of the case to allocate the blame for the "failure" of their children at least partly to the children themselves, however much they may also allocate part of the blame to the teacher and the "school."

Or, alternatively, consider the mixed interests and needs that the different actors bring to their roles in the school system.

The student is focused on evaluated success plus fun plus significant learning plus validation with the peer group.

The teacher is focused on significant learning by the students, plus preservation of his concept of self as a teacher, plus satisfaction of his personality needs in the role, plus validation with his peer group by different criteria than apply to the student peer group.

The parent seeks evaluated success for their children plus significant learning plus continuity between behavior patterns relevant in the school and those relevant in the home, plus evidence of relevance of things learned in school to the lives of the children and to their subsequent careers.

If we add to these diversities, albeit overlapping, of motives and interests that the student, teacher, and parent bring to their roles in the educational system, those which are held by principals, superintendents, and boards of education, teacher-trainers, and others, one can see that though there are some nominally shared central purposes there are also important areas of concern held by some that are of only peripheral interest or even of negative value to the teacher. The parent or child, for instance, who wants the school to devote special attention to a child's particular needs runs athwart the teacher's formal commitment to equal concern and attention to all children.

4. Strains arising from diversity of goals and actors in a democratic system

The problems that arise from attempting to achieve a multiplicity of diverse goals are multiplied once by the presence of diverse actors and several times more when these actors are considered relevant in decisions concerning these goals.

Among the differences most likely to appear are:

1. Disagreements regarding the priority of importance of various goals. "Cognitive learning vs. affective wellbeing" is a summary way to indicate areas of such conflict chronic in educational theory and practice today.

2. Disagreements regarding the allocation of school resources and time and attention to various groups of children. Here the agrument about elite vs. mass education becomes relevant, as does the discussion concerning the propriety of special education for the most disabled children.

3. Disagreements regarding the allocation of teacher competence to various levels of student competence. Once again the "purposes" of the school are here brought into question because of disagreements regarding the special entitlement of the "most talented" children to the best teaching the school can offer. These disagreements inhere in two natural and unavoidable facts: diversity of talent among children and diversity of talent among teachers.

4. Disagreement regarding the methods and approaches to educational goals. Here the disagreement rises out of a very special set of sources: a) the development of a sense of specialized professional skill and competence and a correlative degree of professional pride and identity on the part of teachers and b) the persisting feeling among a community of parents that education is not a specialized skill. Rather it is often held that any layman is as competent as any teacher to decide how schools ought to be run. This conflict over methods and approaches often spills over into areas in which presumably teachers do "know" most about the matter, namely the content of education. Even here, however, there is often dissension.

5. Disagreement regarding the "morals" and "values" taught in the school, or, if not taught deliberately, conveyed and employed by the schools as criteria of student adequacy. Here at least three main sets of actors are involved with their differing definitions of the situation: teachers and school officials, vs. students, vs. parents and the "community at large." Matters are worsened when, as is often the case, these three communities are themselves heterogeneous. And since each of these interest groups is in theory entitled to participate in decisions regarding educational content, procedures, and evaluation, the likelihood of conflict over values and morals is augmented. Here, perhaps even more than with regard to formal educational content, every actor feels

himself quite adequate to the task of making decisions and quite equal to any of the other actors involved. On a petty scale, these disagreements are seen in disputes over proper clothing and hair style. On a larger scale, the conflicts concern such matters as race relations in the school, sexual conduct of the students, and the use of drugs and alcohol. The varying interest of the major groups involved -- their separation by function and generation and their relevant reference groups -- insures a continuing set of disagreements.

When these five types of disagreement are added to the unavoidable disputes about what ought to be taught in the schools, it can be seen that virtually every major facet of education for democratic adult role-playing in a democratic society is a matter with a high potential for dispute.

5. Strains arising from the necessity for continuous evaluation in democratic school systems

In addition to the structural factors just indicated as sources of strain special attention should be called to the contribution of inadequate evaluation of outcomes to the continuity of conflict in democratic school systems.

A corporation can move to reorganize itself if it is inefficient because it learns regularly from its auditors and accountants how "efficient" it has been. So, too, a government can alter its course of policy, both internal and external, because under proper circumstances it is forced by the votes of its constituents, or the tanks of its hostile neighbors, to reconsider its policies. Even a church can get some estimate of how well it is doing, at least nominally, by the size of its attendance and the financial support of its members. But how shall any school or network of schools know how well it is doing?

We shall later devote considerable time to the general problems of evaluation. But now we wish to cite briefly some of the most common problems of educational evaluation and indicate the special dimensions of problems which are added by virtue of the fact that in school systems committed to democracy, with a diversity of goals, ambiguity of content, uncertainty of means, and absence of consensus on criteria, the problems of evaluation are made manifestly and manifoldly more difficult.

1. The clients -- the children to be educated -- are normally considered incompetent to comment on or criticize the operations of the school. They are assumed to be operating from some stance of vested interest in anti-academic, youthful interests that make them biased and incompetent

observers and evaluators. Moreover, except as they come, as they have recently, to take power into their own hands, they have the least formal power or authority in the system of any of the actors. They normally cannot refuse to purchase the goods and services of the "company." They cannot authoritatively locate blame for inadequacies that they sense upon those who have the authority to put the label of success or failure upon them. They cannot withhold relevant resources from the organization without incurring serious reprisals. Since others, notably teachers, have the power to affix the label of failure upon them, and since all other actors in the system conspire to allocate blame for their failure upon the students themselves, the students, though clients, must also be seen as victims of the system and relatively powerless to alter it. Moreover, since blame for "failure" is placed upon their shoulders, the system can continue its chosen path with immunity from impulses to change.

2. A secondary range of clients -- called the parents -- are at a disadvantage with regard to taking effective action in the system by virtue of their systematic ignorance, often self-imposed, with regard to the input and the processes of the system. Such information as they have is usually confined to the very small "bit" called the report cards of their individual children. This tells nothing about the system and its operation but only about the nominal relative score of their children on some arbitrary scales of scores. There are ways by which parents can increase their information about the system, but this requires effort that most parents seem unable and unwilling to exert. However, if they should determine that the system is not operating efficiently -- whatever their source of information or attitude -- they do have recourse of a gross sort. They can vote out the governing board (in the American system), and they can deprive the school of resources, through failing to approve the budget when it comes up for a vote. Occasionally they can, by the same paths, indicate their approval of the system by approving the budget, even with proposed increases. In extreme cases, where the mass media inform them of presumed moral derelictions of one or another teacher or other school official, they can also, as a political public, move toward ousting that official. But their knowledge about, active concern with, and intervention in the content and method of teaching is minimal in most known cases.

As a form of communication useful for change toward greater effectiveness, then, client dissatisfaction and action upon that dissatisfaction is almost trivial.

This impotence of the client and his inactivity regarding his own willbeing within the system presumably designed for his welfare seems to be characteristic of all organizations in which the "clients" are really "captives" of the system. There is evident similarity between students in a school, prisoners in a reformatory, clients in a welfare station, and patients in a free clinic. Because they stand in the relation of impotent client to powerful "servant" or professional, and because they are considered incompetent with regard to the specialized services to which they are exposed, they are infantilized, whether adult in age or not, relative to the professionals; and their "complaints" about the system are seen as reflections of mean spirits and low visions rather than as possibly true reflections of the inadequacies of the system.

3. The goods and services generated and distributed by the system, i. e., the methods and content of education in the schools, are nominally subject to quality control by independent evaluation. But the archetypical mode of evaluation pursued by the schools gives almost no information regarding the quality of its goods and services. For this usual mode of evaluation focuses not upon the adequacy of the content and methods of education but upon the success or the failure of the students in learning them.

Moreover, when there is some scrutiny of methods and content by the agents themselves -- as frequently there is -- it is most often done by the very agents who create the method and content. Hence, little objectivity can be expected. Genuine objectivity regarding methods would require teachers to be critical of themselves. Genuine objectivity regarding content would require teachers to expose themselves anew to the learning process, taking now the role of students themselves. Besides upsetting comfortable habits, exposure to newer methods and new content can be very disturbing because the teacher, now the learner, may "fail" to learn what his new teacher thinks is important. Thus, the conduct at examination time of licensed teachers who are taking courses for advanced credit in teachers' colleges or graduate schools is indistinguishable from that of their own students.

There is further resistance to evaluation be teachers and administrators because evaluation means the possibility of negative results. This is distressful because "success" is crucial to reputation among colleagues and supervisors, and often has serious implications for salary, leaves, and promotions, not to mention the "tenure" syndrome of the early years of teaching.

4. Quality control of methods and content of education is also made extremely difficult by the nature of the teaching situation. How can supervisors -- assuming their competence -- get an adequate sample of the performance of the teachers? This would require careful and prolonger scrutiny in the classroom. Otherwise, one depends on "rumors," or on

the quality of the students as they pass from one teacher to another (here the "badness" of students can always be blamed on the students), or on their performances on standardized tests. Since the latter could be effective and by that token, threatening, great pressures develop to "cram" students for these standardized tests so that they shall show up relatively well as compared with other students in the classes of other teachers. Evaluation through performance on such tests almost guarantees then that there will be no change in the curriculum, since the whole system conspires to preserve the content and methods by striving for excellence in performance rather than by assessing the worth by other criteria. Even this threat of comparative student scores on standardized examinations can be reduced, and evaluation rendered inefficient by breaking up the student body into homogeneous ability of talent groups so that students will perform in accordance with pre-set expectations.

5. Outside evaluations in the form of visitations by certifying boards are almost meaningless insofar as any real scrutiny of method and content of education is concerned. A recent survey of what is done during such evaluative expeditions reveals the amount of dependence on self-reporting in those evaluations, and the lack of critical examination of the schools' conduct. These findings suggest that these so-called evaluations are of only minimal effectiveness, if that, in bringing any impetus for change of method or content to the schools.

6. Even if there were regular and systematic scrutiny of performance by teachers (rather than measuring students' achievements, or in conjunction with such measurement), two structural situations make such scrutiny less effective as an instrument of possible change than it might otherwise be. One is the system of tenure which permits normative incompetence to endure without recourse, since only outrageous violations of codes of professional conduct can bring tenure-rights under reconsideration. The other is the lack of specificity of criteria of adequacy in the teaching role. It is chronic among teachers to insist on the equal validity of different styles of teaching, or to argue that different subjects must be taught in different ways, or to assert the legitimacy of different priorities, such as the importance of concentrating on an elite of the most talented students as against the importance of seeing to it that everyone is taught equally.

A third, related way of defending against evaluation is to query the competence of the evaluators, even though they may be nominal supervisors and presumably have greater formal qualifications of experience. It is argued that these greater "formal" qualifications have no necessary bearing on the competence of the status-superiors to recognize good teaching when they see it.

Because of the lack of specificity in the criteria of adequate teaching, the charge of the incompetence of the evaluators may have some merit. At least it is not demonstrably false. It cannot be falsified at least until common criteria are agreed upon and commonly acceptable measures of the achievement of these criteria have been developed. At the moment, no such criteria or measures are available.

These six difficulties in adequate evaluation of educational outcomes have been offered as sources of potential conflict in the schools. But they must also be seen in reverse. That is, the conflicts in the schools must be seen as circularly and reciprocally contributing to the difficulties of evaluation. Every one of the conflicts listed makes systematic evaluation correlatively and commensurately difficult.

Consider, for instance, the problems for evaluation that arise from the failure to agree upon goals of the enterprise. Or, even when there is agreement on a variety of goals, note the problems raised by the lack of agreement on the relative priority of importance of these goals. The reciprocal reinforcement of conflict and inadequate evaluation can thus be seen as a central feature of modern, democratic educational systems. The lack of evaluation leaves important issues unsettled. And the mootness of important issues, kept that way by lack of evaluation, impedes the development of sound evaluation, precisely because that evaluation cannot be developed until many of the moot issues, such as priority of goals, are settled.

Reprise and Prospect

In a work devoted principally to analyzing structures and strains in democratic educational systems, we should feel compelled to elaborate the preceding materials greatly and to move on, then, to an analysis of modes of reduction of strain and attenuation of stress and solution of conflicts. But our primary aim is not the completion of such a work. Rather we have sought a way to bring into focus the problems concerned with developing a model for the measurement of the effectiveness of educational systems. To do so we have felt it imperative to conceptualize education as an organized enterprise, under varying systems of social organization, and to indicate the points of its vulnerability to stress and strain. Above all, when one comes subsequently to consider how effective various systems have been, the inherently great difficulty in the operations of an educational system when it commits itself to being democratic both in organization and content must be remembered, along with the extraordinary problems presented by the multiplicity of the goals to which democratic systems seem uniquely to commit themselves.

B. FINDINGS AND ANALYSIS

B

3

3. Defining the Goals of Education: Requirements and Ambiguities

A rational policy in education, as elsewhere, involves the choice among competing strategies that are designed to achieve desired ends. The decision-making process and its aftermath proceeds according to an ideal-typical sequence which is as familiar as it is inexorable. Goals are envisioned and sought; means are employed to achieve those goals; there are always disparities between the intended and actual outcomes; the perception of these disparities generates tensions or strains, to which there are reactions, which then have consequences for the next stage of goal-seeking and/or the adoption of new methods. We may illustrate this process by alluding to the problem of "equality of educational opportunity."

1. Educational goal: high-quality schooling should be equally accessible to all categories of American children regardless of their creed, color, national origin, social class, or differences in talent.

2. Correspondence between goal and reality: comparative research in the educational experiences of "standard Americans" and the poor, blacks, Puerto Ricans, and the "dull" reveals that the school system magnifies the inequities of a stratified society by offering a superior education to the children in some groups while systematically denying it to others. Moreover, many such children suffer from environmentally induced disabilities before entering school and throughout their educational careers.

3. Social and individual consequences of the disparities: the individual child experiences anxiety, hostility, and a deflation of self. At the societal level, lack of educational opportunity severely restricts the positive functions of education as a mechanism for recruiting and discovering talent, as an agent for economic growth, as a vehicle for social mobility, and as an instrument for peaceable social change.

4. Research clue: according to some scholars most of the growth or decline in tested intelligence occurs in the preschool years. A child who is the product of an intellectually impoverished environment is severely handicapped by the time he enters first grade.

5. Program: "Operation Headstart" establishes preschool programs as a part of the war against poverty.

6. Evaluation: It is too early to make a definitive judgment, but such programs appear to be valuable. However, there is already sufficient evidence to indicate that they are unlikely to reduce cumulative social and psychological deficits unless they are articulated with subsequent school programs, effectively interpreted to parents and the school systems, and taught by instructors who find gratification in teaching "slow" children.

The function of useful educational theory and research is to make this entire sequence less problematic by rendering ends, means, and their interrelationships more intelligible. It should assist men to decide what goals to pursue, what actions make their attainment more probable, and what are the costs of success or failure. At stake are such questions as: what educational aims do we most cherish? What other goals are we willing to sacrifice in order to achieve them? What are the most efficient means consistent with our values that we might employ to accomplish our objectives? What price are we prepared to pay in scarce resources -- time, energy, and organizational ingenuity -- to achieve our aims? What sectors of society shall bear these necessary costs of attaining our purposes? What shall be the sequence of successive approximations toward ideal goals? What tactics shall we adopt in stimulating consent to our proposals?[1]

As we have not been the first to observe sensible discourse on these matters cannot proceed until we have removed the principal ambiguities which now surround the form and substance of educational goals. American and international evidence suggests that few, if any, societies or school systems formulate their goals with sufficient precision to serve as an adequate guide for action or as standards against which to measure outcomes. The purpose of this chapter, then, is to specify the major formal, political, philosophical, and empirical requirements for the creation of a model of educational goals and to identify the obstacles that impede its attainment.

The Identification of Educational Goals: Verbal vs. Behavioristic Analyses

The most restrictive definition of a goal is an intention to achieve a preferred outcome which is 1) stated in advance, 2) consciously sought, and 3) "realistic" in the sense that it is in principle attainable through purposive social action. A goal thus differs from 1) a fantasy in whose behalf no intervention is thought possible because of intrinsic constraints imposed by nature, social organization, or human limitations; 2) an ideal that is perfectionist in aspiration and indefinitely remote in time; and 3) a *post hoc* declaration claimed after the event.

This pristine conception of the meaning of a goal specifically refers to utterances and rejects as invalid efforts to derive intent from the observation of behavior. As such it may needlessly impoverish our analysis of the "real" aims of American education. Indeed, one influential viewpoint in the social sciences argues that all purely verbal affirmations are suspect and we may be certain that men desire the ends they profess only after we have observed their actions. The credibility of our frequently asserted dedication to equal educational opportunity could hardly be confirmed by comparing the quality of instruction in suburban and slum schools. One need not be a radical behaviorist or a cynic to insist that what men do may be at least as revealing as what they say.

The true nature of educational goals, for example, might be derived from such items as money, time, and personnel actually allocated for specific purposes; outcomes evaluated and honors awarded; the content and balance of the curriculum; the characteristics of teacher preparation, etc. Attention to such indices of intent might help protect us from taking at face value goal statements which are merely ritualistic affirmations of conventional virtue or outright exercises in deceit.

At the same time behavioristic analysis of goals rely entirely on the skills of the observer and suffer from all of the dangers that arise when an original motive is filtered through the perceptions of a middle man. He will always be hard put to make reliable inferences about the complex relationship between intentions and observed behavior. Even leaving aside the preliminary formal problems of avoiding circular reasoning and tautological statements -- whatever people do may be interpreted as the realization of a previously established goal -- we cannot avoid the stubborn empirical truth that human behavior is constrained by limited options. There is seldom any very precise correspondence between desire and possibility. In most cases we cannot be sure whether a particular action-- virtuous or malignant -- occurs because of coercion or choice.

The achievement of the goal of equal educational opportunity, for example, depends on more than philosophical conviction. It probably also requires a prosperous economy that can support education conceived of as both an investment and consumer good; a structure of incentives that includes a demonstrable linkage between educational achievement and the system of social rewards; a normative system that emphasizes the value of secularism, change, achievements, and universalistic standards; a sufficiently stable political structure to encourage institutional and personal investments that have a deferred pay-off, a power structure that can effectively reach and enforce decisions; and a set of mechanisms that reduce actual or potential strains between education and other elements in the institutional complex, especially the kinship and religious system. Since the establishment of some of these conditions is beyond the control of those who profess an allegiance to the goal of equal educational opportunity, we are never justified in assuming out-of-hand that the imperfect realization of a goal is a satisfactory indicator of weak intention. Nor can we be certain that its achievement was a function of anything those nominally responsible did or failed to do. The consequences of social behavior are often neither sought nor foreseen by the participants.

These troublesome considerations suggest that the process of judging the consonance or dissonance between goals and outcomes would be enormously simplified if the former were stated in advance rather than inferred post hoc. If this practice were conventionally adopted we might even discover that much that passes for education proceeds without any clear, well-defined realistic goals. To be sure, even when the aims that move men are not elaborated in precise declarations of purpose the sum of their daily decisions are not without pattern or direction. Nevertheless, the routine requirement to specify goals might reveal that much educational practice is guided mainly by historical legacy, inertia, nostalgia, and considerations of convenience rather than by self-conscious efforts to develop particular competencies or commitments.

Some Systematic Efforts to Specify Educational Goals

The increasing recognition that "success" or "failure" of formal schooling must necessarily be expressed as a relationship between aims and achievement has led to ambitious efforts to develop models designed to detect, record, classify, and establish the connections among educational goals. Taken collectively, their strengths and shortcomings provide a convenient point of departure for our own inquiry into the nature of educational goals.

The companion volumes on elementary and secondary school objectives edited respectively by Nolan Kearney and Will French and the two-volume work on the cognitive and affective domains by Bloom and his colleagues are the most interesting attempts to create useful taxonomies of educational goals.[2] Together, they have made substantial progress in expressing educational goals as observable and measurable behavior outcomes, in establishing priorities, in dealing with developmental processes, and in adapting taxonomic refinements.

Nolan Kearney's Elementary School Objectives is a report submitted by an assemblage of "consultants," mainly academicians who proposed the goals, "critics" who evaluated them in the light of classroom considerations, and a "survey committee" reflecting both theoretical, professional, and practical experience who actively compiled the Report. The aims are prescriptive rather than descriptive: there was no effort made to assess the actual goals of American education but rather to record outcomes preferred by a panel of knowledgeable people.

Their main effort was to classify aims in a three-dimensional grid including domains, behavior categories, and development periods. The domains chosen were essentially broad curricular areas supplemented by analytical constructs designed to measure ethical standards, social-emotional behavior, and physical wellbeing.

Specifically, the domains were 1) health, 2) social and emotional development, 3) ethical standards and values, 4) social relations, 5) the social world, 6) the physical world, 7) aesthetic development, 8) communication, and 9) quantitative relationships. Each domain intersects with a major behavior category -- 1) knowledge and understanding, 2) skill and competence, 3) attitude and interest, and 4) action pattern, and desired outcomes are listed for the "average child" at the end of the third, sixth, and ninth years of schooling.

Kearney exhibits refreshing candor when he concedes many of the major weaknesses of this scheme. It assumes an imaginary child, who progresses at a constant rate of speed, who is comparable in all respects to other average children. Moreover, at the conclusion of the study the consultants agreed that many of the goals were too difficult for average children. The outcomes, themselves, were not ranked in ascending or descending importance. In addition, this volume suffers from a deliberate vagueness occasioned by a failure of nerve. Observe, for example, "The nature of affectional heterosexual feelings is perceived . . . and the individual has acquired adequate and socially acceptable ways of expressing them."[3] All of the hard questions are scrupulously avoided in this formulation.

At the same time many of the goals are expressed with admirable clarity. "He uses a card catalogue, a reader's guide, an almanac, and encyclopedia as needed. He is familiar with the standard library and reference skills."[4] Not "he is beginning to find a sense of restrained adventure in walking in the forests of literary knowledge." Although there are frequent lapses -- "is beginning to exercise some discretion in the value and use of TV and radio programs"[5] -- the work as a whole is a salutary exercise in operational translations of vague concepts.

Will French's Behavioral Goals of Secondary Education relied on research strategy analagous to that of Kearney. The goals in the French work are addressed to three major concerns: 1) growth toward self-realization; 2) growth toward desirable interpersonal relations in small groups; 3) growth toward effective membership in large groups which are cross-classified by four areas of behavioral competence: 1) intellectual growth and development; 2) cultural orientation and integration; 3) maintenance of health; and 4) economic behavior.

In general the French work is subject to the same commendations and strictures as the Kearney volume. There is the same amalgam of tough-minded goals statements -- "spells correctly the words he uses in ordinary written discourse and uses the dictionary if uncertain of spelling" -- and items such as "shows interest in why parents carry life insurance."6

There is little doubt that both the Kearney and the French goal models represent important advances in treating educational objectives. However, they ignore several important desiderata. They confine themselves to statements by experts, neglect the matter of compatibility among goals, and fail to codify the fundamental assumptions on which the goals rest. Kearney and his collaborators make no effort to rank goals in order of their importance while French is content to indicate priorities by placing an asterisk next to those which have been defined as "important" by 95 percent of the expert opinion represented in the study. There is, alas, no evidence that either model has ever been used perhaps because expert prescription does not really reflect the opinion of all the relevant publics.

The model devised by Bloom and his associates differs from its predecessors in its scope and purpose. They are concerned with only two specific psychological domains -- the cognitive and the affective -- and are exclusively interested in goal classification rather than prescription. The scheme contains no substantive proposals; it is a taxonomy that includes among its characteristics the ordering and relating of different kinds of behavior according to a continuum of complexity. The taxonomy in the cognitive domain thus contains six major classes in ascending order: knowledge, comprehension, application, analysis, synthesis, and evaluation. Each class is then likewise sub-classified according to a scale of increasing complexity. Thus, "knowledge" begins with specifics and ends with "theories and structures." The book also contains test items purporting to show the utility of the scheme.

The value of the Bloom scheme, as in the case of any taxonomy, cannot be assessed until there is some large-scale effort to use it for the purposes for which it was designed. This has not yet occurred and as of now its benefits remain problematic. Nevertheless, certain tentative judgments seem permissible. On the positive side the taxonomy is executed with marvelous economy and seeming completeness. To describe the structure of intellect in twenty-five cells is a major tour de force. At the same time the scheme has only a very restricted educational function. Its categories permit the classification of cognitive skills and competencies, but it is substantively vacuous. There is no indication in this taxonomy as to what kinds of information, intellectual products, and knowledge should be arranged hierarchically. It should not be a matter of indifference to schools whether a student develops "theories and structures" about the physical universe or the relative merits of the "T" formation and the single wing as potent offensive weapons for interscholastic football. In short, the taxonomy is of no assistance in the substance of curriculum construction. Moreover, the model is not linked to any particular life goals or role behaviors and it is useful, then, only if we assume that the development of intelligence as such is a desirable outcome regardless of the uses to which it is put and the situations where it is applicable. This is cognitive imperialism writ large and provides another instance of the ways in which methodology creates philosophy. In any case there is no better educational reason to use Bloom than, say, Guilford's structural factors which have the advantage of having been created by factor analysis rather than ad hoc and for which tests already exist.

Kratwol, Bloom, and Musia developed their taxonomy on the affective domain by the same logic as their earlier work. The continuum progresses from a point at which the individual is merely aware of a phenomenon and proceeds through a succession of levels including "responding," "valuing," "conceptualization" to the penultimate stage at which he "organizes" the conceptualizations into a structure which at its point of maximum development comprises his "life outlook." The sub-categories within each of these rubrics are then also arranged hierarchically in ascending order of complexity or conviction. Thus "valuing" is sub-classified "acceptance," "preference," "commitment." The indifference to content which was a serious limitation of the cognitive model is even a greater source of concern in the companion volume. Much of the taxonomy permits us to record progressively finer cognitive discriminations ("conceptualization," "organization") about the affective domain. Although a "commitment" may be distinguished from a "preference" a "commitment" to democratic institutions or totalitarian tyranny would be impartially recorded in the same cell. Since the school is clearly not indifferent to the content of an attitude as well as its intensity the scheme seems to have very limited utility in dealing with a crucial aspect of the affective domain.

It is noteworthy that the only two continuing national inventories of student achievements, Project Talent and the forthcoming National Assessment, do not much advance the state of the art in goal specification or classification. The nature of the outcomes they value nevertheless emerge with various degrees of clarity in general statements, lists of objectives, and measurement instruments.

John Flanagan and his associates in Project Talent are advocates of the "manpower" approach to education. They are primarily concerned with achieving a better "goodness of fit" between the nation's occupational requirements and the available pool of human resources. The absence of

> "definitive knowledge of the capabilities of men is a serious handicap to the nation -- its industries, armed forces, professions, and arts and sciences . . . . The employer who assigns a $10,000-a-year engineer to draft the design specifications that a $7,000-a-year draftsman could do better with greater satisfaction to all involved is wasting money and talent . . . . The guidance counselor who has no tools for predicting whether Johnny will have greater success in accounting or in mechanics falls short of performing a service much needed by Johnny, industry, and the nation."[7]

Accordingly, Project Talent seeks information about seven major areas: 1) the range of available talent qualified for training in various occupational fields; 2) relations among aptitudes, preferences, interests, socio-economic variables, and motivational factors; 3) the consequences of lack of interest and motivation; 4) factors affecting vocational choice; 5) predictors of creativity and productivity; 6) effectiveness of various types of educational experiences; and 7) procedures for realizing individual potentials.

The logic of inquiry which governs Project Talent may be described in a four-step sequence: 1) define the schools' goals in behavioral terms; 2) devise a battery of tests for measuring the child's potential for achieving those goals; 3) develop methods of evaluating instructional materials and practices; and 4) collect data on the subsequent experiences of students with selected characteristics who have been exposed to various curricula and teaching methods.

The final product of this enterprise is a probability statement about the prospects of occupational success. According to Flanagan:

> "To assist the student in planning long-range educational and occupational goals, the counselor would compare the same comprehensive student data in the computer memory with norms based on studies of the experience of students with similar characteristics. For example, if the student indicated he is seriously considering engineering, the counselor can inform him that 80 percent of the boys with this pattern of aptitudes, interests, achievement, and activities who enter college engineering courses graduate."[8]

If Bloom and his colleagues can be chided for their excessive catholicity, for an even-handed indifference to the specific content and direction of educational goals, Flanagan et al. seem unduly preoccupied with a very specific and relatively narrow sector of human experience. For Project Talent the world of work comprises the entire universe which the student will eventually inhabit. The absence of any discernible interest in other social requirements or individual needs apparently reflects a genuine ideological conviction that schools find their justification primarily, if not exclusively, as instruments of vocational preparation. The first report of the project, Design for a Study of American Youth, which lists its long-range aims and research methods, does not make even passing references to the usual commencement pieties about developing aesthetic sensibilities, humane values, or democratic commitments. This omission rather suggests that the neglect of other school-related outcomes has other sources than scholarly specialization.

Flanagan, who himself has contributed to the field of personality measurement, seems to find the psychological dimension salient only insofar as it is associated with occupational performance and job satisfaction. His battery of tests do include a "Student Activities Inventory" which is designed to measure personality but as the author makes clear not because "sociability," "social sensitivity," "vigor," "tidiness," "self-confidence," etc. have any intrinsic value. "It is a well-known fact," write Flanagan et al. "that people of equal ability are not always equally successful in life. Why not? Personality differences, we often say. To probe these differences, Project Talent devised a personality questionnaire, to be included along with the aptitude and ability tests."[9] Manifestly, despite the considerable merits of Project Talent, it will be of little assistance in developing educational goals, except to those who view the United States exclusively as a gigantic labor force and who value schools solely for their capacity to produce efficient and happy workers.

A more expansive perspective on education characterizes the projected national assessment of education as the first organized attempt on a national basis to assess the quality and effectiveness of American schools. The assessment consists of "exercises" (tests, observations, and interviews) which were developed in response to a list of "objectives" which were formulated for ten fields including 1) reading, 2) writing, 3) science, 4) mathematics, 5) social studies, 6) citizenship, 7) art, 8) music, 9) literature, and 10) adult education. The number of objectives for each of the subject areas range from three to seventeen and each is appropriately modified for each of the age groups (9, 13, 17, and 25-35) to be tested.

The directors of the study have taken great pains to indicate what they regard as the significant characteristics of the objectives:

1. They cover only some of the important fields of education.

2. Within a subject-matter field only those objectives are included that are being emphasized in the schools, considered authentic to the

respective disciplines by scholars and teachers and considered desirable by thoughtful laymen.

3. Within the total set there has been a concerted effort to avoid overlap of objectives that could logically be stated under two or more subject fields (e. g. mathematics used in science and citizenship objectives used in various subject matter fields).

4. Although emphasized in the schools, some of the objectives included are not realistic for all students . . .

The above characteristics, then, indicate how these objectives have been developed to serve a specific purpose. They are not a comprehensive set for any school system, not a comprehensive list for any of the subject matter fields included, not a set of standards for individual students, and not a set of objectives to establish future tasks of the nation's schools.[10]

This series of negatives does not exhaust the limitations of the objectives. Characteristics affecting the specificity, range, aspect, and developmental variations of the objectives markedly reduce their elegance and utility.

It is difficult to see how the mandate given to the team responsible for fashioning exercises in the writing area could have been much enlightened by their instructions. The list of goals in its entirety reads as follows: "1) write to communicate adequately in a social situation; 2) write to communicate adequately in a business or vocational situation; 3) write to communicate adequately in school situations; and 4) appreciate the value of writing." One of the four aims classified under Art contains the simple injunction "know about art." The writers of exercises had no cause to complain about curbs on their creativity.

The variations in the number of objectives cited appear somewhat arbitrary and are only partly explained by the intrinsic characteristics of the disciplines. There are only three objectives cited for the entire field of literature (read literature of excellence; become engaged in, find meaning in, and evaluate a work of literature; develop a continuing interest and participate in literature and the literary experience), while the more methodological mathematicians divide their field into ten sub-divisions (e. g. theory of sets, algebra, probability and statistics, concepts of mathematical proof), each one of which typically contains more elaboration than the entire list in literature. It is puzzling why those who enunciate goals in literature are apparently captive to extreme relativist assumptions that "finding meaning in a work of literature" (any work!) is sufficient while mathematicians do not hesitate to require mastery over nearly all the conventional topics of their discipline.

Similar disparities occur in the aspect of learning emphasized by the objectives. The goals require the student either to know (e.g. "understand," "comprehend," "interpret," etc.) or feel (e.g. "appreciate," "respect," "prefer," etc.) or do ("perform," "participate," "use," etc.). There does not seem to be any intrinsic reason why all three types of outcome should not be envisioned for all subjects. Thus it is strange that the list of objectives in mathematics makes no reference to "appreciation" or "joy" or any other feeling state. Surely, mathematics has its own beauty which it is hoped students would learn to cherish. Stranger still, although it is desirable for students of music to "perform" a "musical piece, theme, or figure in any medium," students of art are asked merely to "perceive," "respond," "recognize," "accept" and "know" without ever applying brush to paper or chisel to stone.

The same curious arbitrary quality is also characteristic of the definition of objectives according to age levels. The social studies goal, for example, is "has a reasoned commitment to the values that sustain a free society" and is further specified as follows:

> Age 9: "Respects the views and feelings of other people and can tell why this respect is desirable."
>
> Age 13: "Upholds freedom of speech, the press, religion, and assembly, and can give a reason why he does."
>
> Age 17: "Believes in the value of law and can justify his belief."
>
> Age 25-35: "Believes in open opportunity for advancement and can justify his belief."

The rationale for this developmental sequence is obscure. No characteristic of doctrine (e.g. moral primacy or complexity) or maturity (e.g. increased intellectual power, experience, or responsibility) explains why we should desire or expect that students should become civil libertarians as adolescents and delay until adulthood any allegiance to equal opportunity for all. What prevents a teenager from committing himself to both of these principles?

The churlish tone of some of the comments are not meant to demean the importance of the works reviewed in this section. If the state of the art continues to be poor, it is because the task of defining educational goals, let alone the construction of adequate models, poses extraordinary perplexities. As our subsequent discussion will show, some problems may permanently resist solution.

Some Characteristics of an Adequate Goal Model

A model of educational goals should, of course, satisfy the criteria of any adequate categorical system. It should, for example, be exhaustive. A cognitive model that could not process "develops critical thinking," for instance, would be of very limited utility. Indeed, a classification scheme should ordinarily contain more categories than probable entries, since empty cells are one important measure of the extent to which the model departs from empirical reality.

At the same time, the model should meet the counter-requirement of parsimony. The categories should be sufficiently differentiated to allow for sensitive distinctions but obviously the model must reflect the stubborn fact that some distortion and sacrifice of detail is an inherent feature of all systems of classification. A model is a tool and it will not be used once it becomes too burdensome and unwieldy.

An additional requirement is that the proper location for recording entries can be unambiguously identified. It is important, therefore, that categories should be analytically independent. A model should not include, at the same level of generality, two such categories as "recreation" and "leisure" since presumably the first cannot occur in the absence of the second. Such conceptual overlap markedly reduces the prospect of reliable coding.

This list of desiderata could be extended. An ideal model would provide the opportunity to rank aims according to their frequency and perceived importance and preferably by known internal scales. At still higher levels of sophistication it would permit us to discern to what extent an array of goals are independent or contingent, compatible or antagonistic.

These general specifications for a goal model are, however, not unique to education. The general problems of classification are, in principle, soluble and are actually much less troublesome than the nature of the potential entries -- the actual statements that would be processed in a formally adequate scheme. Now that the Great Debate on "educational aims" that raged in the post-Sputnik era has abated, the casual observer might understandably conclude that all that now awaits solution is finding the means to implement goals held in common. But since the press reports almost daily that one or another community is engaged in bitter conflict on educational policy, it may be that consensus, if it exists, is confined to the highest levels of abstraction and that differences are concealed by the vagueness of existing statements of educational aims. Many of these may be regarded as manifestoes rather than serious proposals or standards for the measurement of educational achievement. In the ensuing sections we shall try to identify the major ambiguities.

Ambiguities of Definition

A goal statement is meaningful when it refers directly or by implication to empirically specifiable and observable outcomes that would allow us to know whether or not a particular aim has been achieved. For example, an aim cited by French et al. fulfills these requirements admirably: "spells correctly the words he uses in ordinary written discourse and uses the dictionary if uncertain of spelling." All of the terms are reducible to extensive definitions, and the behavior itself is accessible to verification by sense experience or theoretical inference.

But what shall we make of sentences such as the following from the President's Commission on National Goals? "The development of the individual and the nation demand that education at every level and in every discipline be strengthened and its effectiveness enhanced." Or how does one deal with the advice that "there must be more and better teachers, enlarged facilities, and changes in curricula and methods."[11] And with the best of will, how is it possible to translate "concern for excellence is imperative" into a program that would fulfill this vague mandate?

A special problem that confounds the interpretation of meaning is that similar phrases have radically different connotations. "Separate education for Negroes" may be espoused by Southern segregationists and black militants but they most assuredly do not share the same goals. The former may view "separate but equal" as a permanent and satisfactory arrangement that assumes white control over black schools, and he will exhibit rather greater enthusiasm for the "separate" than the "equal." Some of the latter may consent to segregated schooling only because there is no other realistic alternative in urban ghettoes; they may demand "community control" of the schools, and their definition of "equality" will include schemes of compensatory education as retribution for past injuries.

Conversely, distortion of meaning can occur when seemingly different phrases have the same connotation and imply the same purpose. Thus, "diversity should be encouraged" and "similarities should be acknowledged" may both be invoked as the rationale for "homogeneous" ability groupings. Infirmities of language may thus exaggerate areas of convergence and conceal crucial differences in existing formulations of educational goals.

Similar observations are reported by a number of the foreign scholars who participated in this study.[12] Van Eyndhoven in reviewing Dutch data showed that the official goals of the major Christian political polities who are committed to religious schooling and the pronouncements of the socialists and liberals, both believers in secular education, are at least in cold print virtually indistinguishable. Even legislation which might be expected to be more precise is actually no more satisfactory. The clause of the Primary Education Act which defines the plan of Basic Education reads as follows:

> "School education, while instructing the children in adequate and useful skills, is to be subservient to the development of their intellectual capacities, to their physical training and their instruction in all Christian and social virtues."

Van Eyndhoven notes that such terms as "adequate and useful" and "social and Christian virtues" are hardly examples of "perfect clarity."

The French, German, and Italian papers also contain illustrations of versatile phraseology. William Taylor in dealing with educational goals from a British perspective reminds us once again that ambiguity may have positive functions:

> "The operational meaninglessness of many goal statements, especially at the level of the whole society, is functional in terms of the need to overlay group and class conflict with a patina of agreement on 'fundamentals.' Such statements sound equally well in the prize day address of the most progressive modernday educator and a latter-day squire. There are, of course, methods by means of which we can 'get back' from generalized statements of aims such as 'the furtherance of democratic values,' 'the encouragement of initiative,' 'flexibility and adaptability,' to what those who state these goals would regard as examples of the kinds of behavior and event they wish to promote. By the use of rating scales, the measurement of concordance, the use of Q sorts and semantic differential techniques, we could, if we wished, find out more precisely what people mean when they make abstract

statements of educational goals. The lack of such studies underlines the ideological purpose of much that is said and written at this level."

We are confronted here with a genuine dilemma for if ambiguity in educational goals performs positive, perhaps even indispensable, social functions then we are obliged to abandon all aspiration for rationality in educational planning. We shall never be able to make rational choices among competing alternatives until educational goals are precisely formulated and the putative relationships between ends and means have been specified and the actual correspondence between avowed purposes and observable outcomes have been adequately measured.

Ambiguities of Incomplete Reference

Every goal statement, if it is to be genuinely useful, must specify clearly whether it is applicable 1) to all youth or only to some, 2) for all roles or only to some, 3) all the time or only part of the time.

All or Some

Influential segments of all nations represented in this study are currently committed to expanded educational opportunity. The United States has clearly the longest and perhaps the most serious allegiance to the most noble and ambitious of all educational goals -- equal educational opportunity. Nevertheless, there is reason to believe that ancient rivalries between egalitarians and elitists still persist.

The egalitarian creed expresses an equal concern for equal quality education without reference to group or categorical membership and regardless of differences in talent. To the question of who shall be educated it answers "everybody." For how long? "A lifetime." At what level of excellence? "The highest." Elitists believe in universal education but they urge a disproportionate expenditure of resources and concern on the gifted few. Thus, the omission of the phrase "regardless of differences in talent" may reveal a strong if unexpressed conviction about one of the most fundamental issues in current educational policy.

In general, it is dangerous to assume that in the absence of an assertion to the contrary a stated objective refers to the entire population. James Conant is an eloquent advocate of the comprehensive school and he has defended it "on social and political grounds as an instrument of democracy, a way of mitigating the social stratification of society." Yet he confesses that he "said not a word to indicate that certain schools I visited were comprehensive only insofar as white youth were concerned."[13] Similarly, Education for All American Youth, a book much admired by the liberal establishment, likewise favored the comprehensive school but remained discreetly silent about the exclusion of Negroes. This constitutes an interesting demonstration that even the seemingly unequivocal word "all" can be burdened with ambiguity.

For over a decade there has been increasing recognition that "equal education does not mean identical education." Despite almost universal consent to a proposition that has become a cliche, most goal statements have only now begun to identify distinctive educational aims on the basis of race, sex, ethnicity, and social class level. These taboos had apparently been observed to avoid injuring the sensibilities of "minorities" who like women, the poor, Negroes, and Puerto Ricans have been the objects of educational prejudice. The putative victims, themselves, have increasingly indicated their desire to have their special needs recognized.

Spokesmen for the disadvantaged have questioned the desirability of imparting middle-class values to all children, some women regard an "academic" education as peripherally relevant to their adult lives, and significant segments of the black community wish to emphasize its distinctiveness rather than its points of common identity with white society. Unless the issues of group differences -- not to mention personality variations -- are explicitly confronted in statements of educational aims, an outmoded conception of equality will deflect our attention from the differential impact of school practices in a diverse population.

All Roles or Only Some

Education is always in some degree oriented to the future. One of its primary goals is to develop in learners the competencies and commitments that are required for participation in adult life. The social system in which they will participate consists of roles and their interrelationships. Role occupants grow old, weary, and die so that every society must so to speak "manufacture" replacement parts. Students are at minimum taught explicitly or through subtle cues age, sex, and stratum appropriate roles as these are acted out in such major life tasks as work, citizenship, love, friendship, and play. They are also instructed in the degrees of freedom available to them in deviating from role definitions. The variations in compulsion are conventionally expressed in sociology as "prescribed," "proscribed," "preferred," and "permitted."

The responsibility for producing adequate role performances is shared by many institutions in industrial societies but by common consent it falls within the special province of the family and school. The family is in one sense the "primary" institution since it precedes formal educational systems in historical time. It is also the first, most diffuse, and perhaps the most enduring influence exerted on a child.

There are, however, at least two features of Western industrialized nations that entail socialization processes that are beyond both the expertise and the jurisdiction of the family. The young must now be prepared for a society characterized by increased 1) social differentiation which requires extraordinarily varied and complex role performances and 2) social solidarity which requires loyalties to larger units beyond the kin group. These general tasks, as well as others, are increasingly entrusted to the schools.

There are various ways to classify role-related educational goals. These may be reduced to 1) specifying functional tasks (a la the Lynds' "making a home," "training the young," "using leisure," etc.); 2) identifying appropriate institutional areas (e.g. economic, religious, kinship, etc.); and 3) indicating nonrole specific attitudes that are presumably relevant for a variety of tasks and areas (e.g. health, emotional maturity, etc.).

Initially we preferred a scheme comprising seven major institutional roles and nonrole-specific characteristics as follows:

1. economic - including both production or service role and consumption role;

2. political - including the varying forms of possible participation in the political process from holding office to membership in party to voting;

3. primary group membership roles - including both those connected to kinship and reproductive functions and structures, and those found in non-kin structures and relationships (e.g. friendship, love, etc.);

4. socialization and/or educational activities - including everything concerned with the activities of the person as a socializer or educator of the young on the one hand, and his own involvement in education beyond the cutoff point used here to locate educational outputs;

5. activities connected with the absorption, understanding and various participations in the cultural heritage of the society - including science, art, philosophy, religion, play, and recreation;

6. orientations to self and others - including intrapersonal, interpersonal, and intergroup attitudes, where otherwise not classifiable as specific to another role area, and intended to refer to the generalized person;

7. physical wellbeing - including the prevention of disease and the maintenance of health.

As several of our European colleagues have forcibly brought to our attention this scheme is vulnerable on a number of counts which are discussed in greater detail in later sections. But whatever the merits of particular modes of classification it seems apparent that the school can and does prepare students for participation in a wide variety of adult roles whether or not this is its specific intention. In point of fact there has been more self-consciousness about this issue than about most problems of educational philosophy. Indeed much of the controversy surrounding American educational aims has centered on the primacy of the activities that we have identified as "absorption, understanding, and various participations in the cultural heritage of the society" as distinguished from other forms of participation in social life.

The nature of the continuing dispute may be traced by referring to a sequence of goal statements by the influential National Educational Association. In 1892, the NEA Committee on Secondary School Studies, composed mainly of representatives of higher education and chaired by President Eliot of Harvard, stated that the purpose of education was to train and discipline the mind.[14] But by 1918 the single most influential of all statements of educational goals, "the seven cardinal principles" fashioned by the NEA Commission on Reorganization of Secondary Education, adopted a more catholic view.[15] The principles were designed to "develop in each individual the knowledge, interests, ideals, habits, and powers whereby he will find his place and use that place to shape both himself and society toward ever nobler ends."[16] This formulation has a certain resemblance to doctrines advanced by "progressive" educators but the actual principles owed their form to a more conservative thinker. Herbert Spencer had written that the life tasks of people, not an arbitrary set of eternal truths, should determine the nature of educational goals. He noted five such areas: "1) self-preservation, 2) securing the necessaries of life, 3) rearing and discipline of offspring, 4) the maintenance of proper social and political relations, and 5) the activities which make up the leisure part of life, devoted to the qualification of the tastes and feelings."[17]

The seven "principles" are an extension and modification of the Spencerian "activities." Education was to be measured by its contribution to "1) health, 2) command of the fundamental processes, 3) worthy home membership, 4) vocation, 5) citizenship, 6) worthy use of leisure, and 7) ethical character." Except for the fundamental processes which refer to elementary cognitive skills such as "reading, writing, arthmetical computations, and the elements of oral and written expression" there was otherwise no reference to the life of the mind. An admonition to teachers of English warns them that "only so much theory should be taught at any one time as will show results in practice."[18]

The Seven Cardinal Principles survived virtually unchallenged as the prototypical document on educational goals until the Great Debate on the schools during the post-Sputnik decade. Thus a 1938 NEA report listed "self-realization," "social relationships," "economic efficiency," and "civic responsibility" as major aims.[19] However, in 1961 the NEA Educational Policies Commission declared that the "central" although not "exclusive" purpose of the school is "the development of every student's rational powers."[20] All of the other aims summarized by the Cardinal Principles are subordinate to the capacity to think. Health, for example, depends upon a "reasoned awareness of the value of mental and physical fitness and of the means by which it may be developed and maintained."[21] Similarly "more than ever before, and for an increasing proportion of the population, vocational competence requires developed rational capacities."[22]

The distance between the 1918 and 1961 goal statements is not as wide as it first appears. The latter declaration did not formally reject non-cognitive objectives; it merely identified the "fundamental processes" as the prime means by which all other aims could be achieved. A subsequent report, Schools for the Sixties published by the NEA in 1963, confirms the impression that at least at the abstract level the "cognitive vs. frills" conflict had lost the capacity to produce fiery polemics.[23] This volume also assigns the "highest priority" to the cognitive sphere -- "reading, composition, listening, speaking, (both native and foreign languages), and computation, . . . ways of creative and disciplined thinking, including methods of inquiry and application of knowledge" -- but it also lists the cultivation of aesthetic taste and the promotion of health as important aims of the school. Moreover, instructional goals take first priority because the curriculum is by definition a "distinctive" school responsibility. But according to Schools for the Sixties the school also shares with other social agencies jurisdiction over such aims as the development of values and ideals, social and civic competence and vocational preparation.

The apparent reconciliation of perspectives should nevertheless be viewed with some skepticism. Common observation suggests that most school systems make invidious comparisons among children on the basis of their presumed "native ability," periodically administer aptitude and achievement tests which measure academic aptitudes and proficiencies, award letter grades to students only in subject matter areas, issue annual reports boasting about the proportion of their graduates who are admitted to elite colleges, reserve their most prestigious rewards for students who excel intellectually and otherwise testify that they measure accomplishments in the cognitive domain above all other educational outcomes. Psychological development, social attitudes, and "citizenship skills" are not systematically evaluated and often not noticed unless they obstruct "classroom management," that is to say, interfere with "the training of the mind."

The emphasis on cognitive achievement seems to extend also to learning modalities as well as domains. In common usage, the educated man is one who knows rather than one who feels or does. The educator's fear of "value indoctrination" and the frequent warnings against "manipulating personality" have many sources, most of them doubtless noble, but here, as elsewhere, freedom and autonomy are sometimes granted because of indifference rather than philosophical conviction.

The widespread acceptance of the progressivist dictum that "children learn by doing" does not mean that active "participation," "application," and "rehearsing" enjoy equal status with cognitive outcomes. "Doing" so conceived is an instrumental device which is subordinate to the objective of imparting information and understanding. Our impression is that mock parliamentary assemblies, for example, are regarded as vivid pedagogic means of helping students understand the legislative branch rather than as direct preparation for citizenship. If rehearsing for future participation in community affairs were taken seriously as an authentic educational concern, the student government might not be conceived as a superimposition on the real business of instruction. Indeed, if preparation for a full range of adult roles were an important curricular objective, it is difficult to see why, except for an occasional field trip, instruction takes place almost wholly in the insular world of the school building.

We may anticipate that the controversial question of what constitutes "education for life," now temporarily dormant in the United States, will become increasing salient in other Western industrialized nations. As greater numbers of children spend more years in school the curriculum at each age level is designed largely to render the students eligible for the next stage of schooling. Education thus tends to become a semi-autonomous enterprise governed by its own norms and values which are based primarily on respect for achievement in the cognitive domain. At the same time the entrance into the educational system of large numbers of less academically talented children whose parents tend to view education as an instrumental rather than an intrinsic good, exerts pressures to broaden the purposes of the school beyond preparation of cultural consumers. The tension between cognitive and non-cognitive objectives is a problem inherent in any complex society which sponsors a system of mass education. Any goal model which does not allow us to record these strains will lose much of its analytical power.

Now, Later, or Always

Educational goals may be differentially relevant for individuals at various stages of the life cycle. This truism has been recognized in some of the works surveyed in earlier sections of this chapter. However, most goal statements are seldom very fastidious in specifying whether a particular aim is equally suitable for young children, adolescents, mature adults, or the aged. It is obvious that some objectives must be viewed developmentally and progressively modified; the child who is originally urged to "achieve gradual independence from his family of orientation" may also be asked to serve as the chief source of financial and emotional support for his parents in their declining years. Other goals, once achieved ("the child should develop an appreciation for the cultural contributions of other peoples"), should presumably persist for an entire lifetime. The failure to make distinctions between now, later, and always has had the incidental effect of inhibiting social and psychological research on the stability and continuity of behavioral attributes and has made it more difficult to discern whether desired outcomes have any genuine prospect of persisting during the adult years.

There is now no reliable information on the connection between study and behavior and/or values, once sufficient time has elapsed to allow for the intervention of contravening influences. It is conceivable that the methodological hazards of conducting researches that would yield such knowledge are insuperable. If this is indeed the case we shall have no alternative except to reconcile ourselves to the fact that evidence relating to the achievement of some kinds of preferred outcomes can only be established either by faith or theoretical inference. The careful specification of the temporal provenance of goals may well have the incidental function of defining some of the limits of what we may ever hope to know about the relationship between goals and outcomes.

Ambiguities of Relationship Between Goals

Priority Rankings

The statement of school objectives are often coextensive with the full range of desired human behavior. French _et al._ include among their desired competencies "he helps when necessary to eliminate insects and vermin which tend to carry germs"; she "wears with growing self-assurance foundation garments and clothing properly styled for the maturing figure"; he "stands for and defends the right of each individual to worship God in his own way or refrain from religious affiliations or beliefs." The failure to distinguish important from trivial objectives could make nonsense of the entire enterprise.

An even more troublesome problem in establishing priorities is how to chose between two or more potentially incompatible goals both of which have some claim to legitimacy. A common strategem is to state each with sufficient vagueness to obscure their more extended implications. Thus, in the goals of the NEA Educational Policies Commission we learn that the "educated citizen is sensitive to the disparities of human circumstance," "he can work and play with others and he puts human relationships first." But he is also, in the language of the report, "economically literate."[24] Suppose economic literacy should lead him to the conclusion that financial success is sometimes incompatible with Christian ethics? Shall he then adopt the ethics of the counting house at work and strive for moral perfection elsewhere or live by identical standards in all sectors of his life? Where and when may he chose to guide his actions by the norms of transaction, bargain, and mutuality of reward in preference to "spontaneity," "openness," and "totality" or must he always behave as if he has a friend at Chase Manhattan? The goal statements do not tell us.

The educational philosopher James McClellan has recently argued in essence that goal priorities can be established only under very special circumstances and never when we must chose between instrumental and intrinsic values. He specifically rejects a schema which he attributes to James Bryce Conant and which he summarizes as follows:

> Let us summarize these remarks in the form to be found in Conant's arguments at that point in his career as educational theorist when he was most seriously seeking a synoptic view of the System he had helped to build; he held to the simple schema S which may be diagrammed as follows:
>
> S
>
> Educational Practice   Ordered System of Values V
>
> C
>
> where C, the connection between the educational practice and V, is: the practice is the value sought. I have argued that S is a logically feasible argument only if V contains but one kind of value at a time, and that it breaks down if you try to mix intrinsic values and instrumental values in the same ordering system. I submit that Conant's own work demonstrates the futility of this attempt when he is forced to say that the justification for teaching the humanities is the contribution this study makes to the mental health and consequent political stability of our citizenry.

([Education for a Divided World], p. 92) That sort of patent nonsense is unavoidable unless one is willing really to restrict his attention just to instrumental values, and Conant is much too much a humane gentleman for that.

Thus S will work so long as only the instrumental value of schooling is considered. Conant's scheme of 1948 was fairly simple: the preservation of the American way of life in the face of its myriad of enemies. And for him the American way of life had three related sub-systems: a) an honest, decentralized, federal political system, b) a laissez-faire-as-possible economic system, and c) a fluid, complex social system, probably in that order. If we can establish some such ordering of non-educational values, we can in principle order activities in school by their contribution to that system of values, providing a basis for intelligible discourse at least. The trouble is, of course, that once we put distinctively educational values -- intrinsic, for-themselves values -- in the schema, we have incommensurables, and the whole notion of ordering is lost. There is simply no single standard by which to measure the relative value of mastery of a complex proof in pure geometry on the one hand and learning the routines necessary to perform adequately as a service station attendant on the other.

Both of these activities are values, the first intrinsic, the second instrumental. There are many things one can say about both these values, e.g., that both are important in a total society, that both take some diligence and attention, that there should be no invidious distinctions between individuals because of particular talent or preference for one over the other, etc. A well-known John W. Gardner made a big deal of saying those things one may say about both sorts of values. But none of that will help S. S depends on the logical possibility of comparing various school programs by their contribution to an ordered scheme of values. But mastering a proof in pure geometry, sharing the unity of sense, feeling, and form of "Es war ein Konig in Thule," and values like that cannot be ordered on the same scale with learning the skills necessary to get a job, and to vote, and other values like that. Hence, S will not do when complex problems of educational policy arise, where "complex" means requiring decisions involving incommensurable values.[25]

These are persuasive and powerful arguments but their burden is only that there is no logically acceptable way by which it is possible to order what McClellan calls "incommensurable values." There do exist, however, other less rigorous standards for making choices, although they are, to be sure, frequently philosophically naive. Several come to mind. Priorities could be established on the basis of 1) arbitrary preferences, 2) theories of individual and social welfare, 3) consensus of "qualified" observers, 4) intensity of feeling, 5) the number of domains for which an aim is relevant, 6) the length of time for which the projected outcome is intended, 7) the perceived urgency of immediate implementation, and 8) ideological derivation.

None of these is wholly adequate. The failure of philosophy to produce a mature axiology means that we have no satisfactory way to deal with the value problem but it does not relieve us of the necessity to make choices. And it is never enough to catalogue aims under appropriate categories unless we also indicate whether we wish to assign them equal weights. Since it is not possible to do everything at once, and since some aims are in some sense subordinate to others, the failure of most declarations of goals to include priority guides constitutes a very severe limitation on their utility.

The Special Problem of Society and the Individual

Walter Robinson Smith, one of the first American educational sociologists, has argued that discussions on education until comparatively recently exhibited a pronounced individualistic bias, that indeed "the whole literature of education, prior to the last quarter of the nineteenth century, may be searched in vain for any clear statement of the social purposes of a scheme of formal education."[26] An examination of definitions of education offered by prominent thinkers lends credence to this contention:[27]

Plato - "Education consists of giving to the body and the soul all the perfection of which they are susceptible."

Aristotle - "The true aim of education is the attainment of happiness through perfect virtue."

Comenius - "Education is the development of the whole man."

Locke - "The attainment of a sound mind in a sound body is the end of education."

Pestalozzi - "Education means a natural, progressive, and systematic development of all the powers."

Froebel - "The object of an education is the realization of a faithful, pure, inviolate, and hence holy life."

There is abundant evidence to the contrary. Well before the last quadrant of the nineteenth century, some writers had viewed education in social terms. Socrates can be interpreted in this fashion -- "The aim of education is to dispel error and to discover truth" -- and Milton's social orientation is clear: "I call therefore a complete and generous education, that which fits a man to perform justly, skillfully, and magnanimously all the offices, both private and public, of peace and war." In any case, most definitions of American education formulated in the recent past include a social component and many place their primary emphases on collective, rather than egocentric, purposes. Observe, for example,

Ward - "Education means the universal distribution of extant knowledge."

N. M. Butler - "Education means a gradual adjustment to the spiritual possessions of the race, with a view to realizing one's own potentialities and to assisting in carrying forward that complex of ideas, acts, and institutions which we call civilization."

J. J. Findlay - "Defined in its broadest terms, education is no more and no less than the provision that mankind has to make for the progress of the species to which he belongs."

J. Dewey - "Education signifies the sum total of processes by means of which a community or social group, whether small or large, transmits its acquired power and aims with a view to securing its own continuous existence and growth."

The President's Commission on National Goals, when it appeared in 1960, reflected the American consensus that education must serve both personal and social aims. "Our deepest convictions," wrote John W. Gardner, "impel us to foster individual fulfillment. We wish each one to achieve the promise that is in him." And he added, "Ultimately, education serves all our purposes -- liberty, justice, and all our other aims . . . ."[28]

Gardner's remarks reflect a notable tendency in American discussions on education to assume that the interests of society and the individual coincide. The development of "citizenship skills," for example, is customarily included as a worthy goal on the dual grounds that a student so trained will be equipped for the good life even as society collectively benefits from the exercise of his civic virtue. The educational literature seldom defers to the tragic view of life. It does not ordinarily recognize any strain between the imperious yearnings of the individual and the brute facts of social requirements and constraints. This is in many ways a remarkable phenomenon for there are strong contrary traditions which convey counter messages. The entire thrust of classical liberalism was devoted to protecting a sphere of inviolable rights against an intrusive state. American-Soviet ideological differences have been interpreted in similar terms. Freudian theory has identified civilization as a necessary but inevitable creator of discontent, a popular literary tradition whose archetypical symbols are Huxley and Orwell has identified society as the enemy. The "hippies" are only the most recent and dramatic illustration of the effort to seek salvation far from the madding crowd.

The failure to confront the society-individual dilemma has seriously impaired sensible discussion on a number of issues which have recently engaged the attention of educators among them sex education and the development of "autonomous personalities."

The most notable characteristic of stated aims on sex education is that they are relatively content-free. Teachers are advised to answer all questions, be natural, avoid self-consciousness; they are told everything except what most concerns them, i.e., what constitutes desirable sexual behavior for young people. The students themselves are often clear enough. An increasingly popular sexual ethic fails to perceive any moral significance in the physical act of sex aside from the quality of the relationship that unites the partners. There can be no sensible analysis of this contention without considering the impact of sexual freedom on monogamous marriage and arrangements for child-rearing. Can we maintain the institution of private children while permitting sex both in and outside of marriage to become a matter of private definition? In answering this question for ourselves and students we shall need to mediate between the individual's desire for adventure and novelty and society's concern for a stable family system in which each marital pair is responsible for its own progeny. We are reminded once again that every theory of education is simultaneously a theory of society.

Another case in point is the aim of teaching people how to become non-conformists, self-directed, and unresponsive to the middle-class norms and the bureaucratic ethos that pervades American life. But the fact is that industrial nations require citizens who are work-oriented, responsible, and masters of tangible skills, and the emphasis on such competencies tends to generate stable lifestyles. It is an open question whether it is possible to organize society so that it is at once prosperous and habitable, that is to say, humane, spontaneous, and free. There may be no inherent contradiction between individual and group existence but many men have thought so and it is a disservice to more precise formulations of educational goals when so fundamental an issue is solved by the simple expedient of ignoring it.

Sequential Derivations

Most statements of educational goals are not ordinarily accompanied by an adequately detailed geneology of their intellectual ancestry. The line of descent, if traced at all, usually extends only so far as to one of the warring schools of educational philosophy -- perennialism, essentialism, progressivism, reconstructionism, and more recently radical humanism.[29]

The most prominent representatives of the perennialist tradition are Robert Hutchins and Mortimer Adler. The basic assumptions of perennialism may be succinctly summarized by Adler's assertion that "man is a rational animal, constant in nature throughout history"[30] and by Hutchins' syllogistic description of the educational process: "Education implies teaching. Teaching implies knowledge. Knowledge is truth. The truth is everywhere the same. Hence, education should be everywhere the same."[31] Societies rise and fall, feudalism yields to industrialism, the spear is replaced by the guided missile, but human nature and the eternal verities endure. The distinctive property of man is reason and the school trains the mind by exposing it to the accumulated wisdom of the past. The school has timeless concerns and does not include among its aims vocational preparation, social reform, or any other purpose not specifically devoted to the training of the intellect.

The progressivist followers of John Dewey entertain an entirely different world view. It is a perspective that rejects absolutes, that finds change and flux everywhere, that regards past solutions as treacherous guides to novel and unpredictable events. Education, then, does not consist of transmitting a durable legacy, but rather equipping children with the disciplined intelligence to deal with change. And since children differ in their genetic and social characteristics so must their education. All young people are capable of individual growth, and when proper provisions are made for individual differences they can all benefit from education at some level. The progressivist task entails preparing all children to perform a variety of complex social roles of problematic stability.

The conflict between perennialists and essentialists or between progressivists and reconstructionists is devoid of real bitterness. Essentialists are less persuaded about the importance of conveying eternal truths than the perennialists but they too regard the inculcation of prescribed subject matter in an intellectually demanding fashion as the primary purpose of the school. Reconstructionism may be less beguiled by "critical thinking" and "disciplined intelligence" than the progressivists and prefer more positive efforts to build a new society based on the methods of social science and the ideals of socialism but reconstructionists, like all millennialists, have often been obliged to enter into coalitions with their nearest intellectual neighbors.

Paul Goodman, Edgar Friedenberg, and others have recently championed a conception of the school which while not systematically developed has figured prominently in educational discussions. These thinkers have advanced a radical critique of education which holds that the schools now reflect the meaninglessness, conformity, and "manipulative" ethos of a broader society that is dominated by sterile middle-class values. Instead, the aim of education should be the development of autonomous and creative personalities that will maintain their "fidelity" to self and their fellow men. Education worthy of its name might assist children to survive in a corrupt society but it would not assert the moral superiority of the "marketing orientation" that is presumably necessary for"success"in school and community. Since Goodman and Friedenberg regard most professional educators as enemies of spontaneous and open creative thought and lifestyles, and they offer few viable programmatic alternatives, their ruminations have thus far aroused more enthusiasm among unattached intellectuals than practicing schoolmen.

If we examine the structure of the argument rather than the content of these philosophies, we discover that each proceeds according to an implicit sequential paradigm beginning with master values, proceeding through master empirical images, master social (or individual) goals, and then finally educational goals. Some values are preferred as ultimate or "primitive" notions, a particular image of man or society promises that there are no intrinsic barriers in the empirical world that need frustrate the achievement of these values, a tangible goal is defined which serves the value, which in turn implies a corresponding educational aim. The entire process may be illustrated by a hypothetical sequence chain beginning with the master value "freedom":

Master values	- Freedom is preferable to slavery.
Derived values	- Political democracy is more desirable than totalitarianisms of the right or left.
Master empirical images	- A system of countervailing power will guarantee the preservation of democratic institutions.
Derived empirical assumptions	- A system of checks and balances among the executive, legislative, and judiciary is essential for the maintenance of democratic institutions.

Master social goal - The preservation of political democracy.

Derived goal - The preservation of systems of countervailing power.

Master educational goal - The development of competencies and commitments required to maintain political democracy.

Derived educational goal - The development in students of a willingness to recognize "Supreme Court" decisions as the "law of the land" even if they are personally offensive.

These sequential derivations are rarely acknowledged or made explicit and therefore we cannot discover whether all the intermediate points on a "scale of ultimacy" are either logically linked or consistent with current knowledge. We do not know, in short, if the goals are part of a coherent intellectual system.

In order to satisfy ourselves on this score we would require categories capable of cross-cutting various analytical levels, i.e. values, empirical images, goals, etc. A common practice is to rely on the treacherous left-right continuum. We could, for example, identify some paired couplets the first term of which was arbitrarily designated as "conservative" and the second as "liberal" and try to discover to what extent a preference for one or the other in any given pair was consistently associated with similar choices for all of the items. Thus *social values* might be represented by "order-change," "station-equality," "consent-participation," "law-justice," "struggle-security," etc.; *behavior science* by "eugenics-euthentics," "trait-Gestalt," "superego-ego," "static-dynamic," "punishment-reward," etc.; and *educational theory* by "perennialism-reconstructionism," "elite education-mass education," "authoritarian-democratic," and "grading-evaluation."

Our guess is that it is possible to identify a liberal or conservative syndrome; there is a certain conceptual kinship between "station," "static," "perennialism," and "elite education" as there is also between "equality," "dynamic," "reconstructionism," and "mass education." Nevertheless, the terms "liberal" and "conservative" are so slippery and their referents are so imprecise and controversial that many scholars have been led to despair of their utility, even as heuristic devices.

The pattern-variable scheme developed by Talcott Parsons provides another alternative for attaining some measure of coherence in sequential derivations. Parsons contends that any system or subsystem level, social action involves choices among a range of polar alternatives. In addition to deciding between competing loyalties to self or collectivity the actor is confronted by four other dilemmas which Parsons defines as follows:

1. Affectivity-Neutrality. Should the actor react to the situation by immediate emotional involvement, or shall he respond with an uncommitted, disciplined attitude?

2. Universalism-Particularism. Should the actor be governed by standards that are equally applicable to all, or shall he be guided by the particular relationship that he has with a person, role, or value?

3. Specificity-Diffuseness. Should the actor be oriented toward a broad and ill-defined range of attributes, or shall he be concerned with only one aspect of an "action object"?

4. Ascription-Achievement. Should the actor treat the object of action on the basis of his qualities -- what he is -- or should primacy be given to his actual performances?

The pattern-variable scheme does not include any formal rules of classification and as such is vulnerable to problems of reliability among observers. The decision to assign a choice to the category defined by one or another of the end points on the continuum, or for that matter the determinance as to which pattern is relevant, must necessarily be based on such loose criteria as "plausibility," "common sense," and "insight." Nevertheless, as the following diagrams indicate the pattern variables may be a useful means for imposing unity on otherwise unconnected entries at several interrelated levels of analysis: 1) social values, 2) curriculum content, and 3) control of classroom by the teacher.

	I Neutrality	II Affectivity
Goal	Analysis	Involvement
Curricular Perspective	Emphasis on objectivity and detachment	Emphasis on subjectivity and emotion
Teacher Orientation	Therapist as role model	Parent as role model
	Universalism	Particularism
Goal	Uniformity	Diversity
Curricular Perspective	All study the same things	Variety of options
Teacher Orientation	Teacher's relationship to students same for all	Students treated according to their special characteristics

	I Specificity	II Diffuseness
Goal	Knowledge	Breadth
Curricular Perspective	Exclusive emphasis on cognitive area	Pluralistic emphasis embracing cognitive achievements and including also values and personality
Teacher Orientation	Guide students in school subjects	Guide students in a wide range of life problems
	Ascription	Achievement
Goal	Stability	Change
Curricular Perspective	Inculcation of revealed wisdom	Development of instrumental skills
Teacher Orientation	Children rewarded according to self-fulfilling prophecies linked to their social status	Children rewarded for their performances

Column I (neutrality, universalism, specificity, ascription) and Column II (affectivity, particularism, diffuseness, achievement) each constitute a more or less coherent ideal-typical system of allegiances to specified preferred outcomes, educational theories, and pedagogic practices. This mode of analysis can be extended to the selection of means as well as goals. The pattern variables may be helpful in organizing choices with respect to the proper time, type, scope, and source of action.

	I Neutrality	II Affectivity
Time of action	Adoption of long-range rational strategy	Intermittent and sporadic response to dramatic crises
	Universalism	Particularism
Type of action	Same for all situations having same elements	Varies on basis of fluctuating reaction to other actors and situations
	Specificity	Diffuseness
Scope of action	Concentration on a single issue	Concentration on a wide spectrum of issues
	Ascribed	Achieved
Source of action	Action based on traditional tactics and strategy and determined by ritualistic considerations	Action based on experimental tactics and strategy and determined by pragmatic considerations

The instrumental means summarized in Columns I and II seem, on intuitive grounds, to be extensions of their counterparts in the goal model. This entire exercise is, of course, merely illustrative of the type of unifying concepts that could be imposed on the entire sequence chain beginning with master values, and extending through empirical images to social and educational goals and means. The major purpose of this somewhat lengthy detour is to emphasize that it is urgent, and to a limited degree possible, to examine the entire structure of belief, value, and knowledge that sustains an educational goal. Does it follow, to refer to only one cell in our truncated diagrammatic presentation, that a generalized preference for "stability" necessarily presupposes a curriculum anchored in the past, and teachers who respond to students primarily on the basis of their group and categorical memberships, or can the master value be reconciled with all manner of educational philosophy and mode of classroom interaction? Are the imputed relationships linked by fiat or intuition or perchance by reason and evidence? These questions cannot be explored, much less answered, in the absence of a reasonably detailed and comprehensive account of the assumptions on which educational goals rest.

The aim of this chapter was to provide some indication of the formidable range of problems that must be solved prior to the construction of an adequate goal model. The entire process is further complicated, as we have indicated, by the number of interest groups -- educators, parents, students, etc. -- that have a legitimate stake in the education of children with the result that we shall need not one but rather multiple goal models each representing the perspective of a distinctive public. Clearly then, much remains to be done before stated aims can be used as a measure against which to weigh outcomes -- and as we shall see the requirements for establishing the claims of education are no less severe.

REFERENCES

1. Parts of the following two chapters also appear in several essays by Marvin Bressler which appeared during the course of the present inquiry. These include:

   "Sociology and Collegiate General Education" in Paul F. Lazarsfeld, William H. Sewell, Harold L. Wilensky, (eds.), *The Uses of Sociology*, New York, Basic Books, 1967, pp. 45-62.

   "The Potential Public Uses of the Behavior Sciences" in Frederick T. C. Yu, *Behavioral Sciences and the Mass Media*, 1967, pp. 3-27.

   Defining Profitable Areas for Research in the Sociology of Education, Office of Education, 1968, Project No.F-046, 93 pp.

2. Nolan C. Kearney (ed.), *Elementary School Objectives*, New York: Russell Sage Foundation, 1953; Will French (ed.), *Behavioral Goals of American Education*, New York: Russell Sage Foundation, 1957, p. 97; Benjamin S. Bloom, et. al., *Taxonomy of Educational Objectives, Handbook I: Cognitive Domain*, New York: David McKay Co., Inc., 1956; David R. Krathwohl, Benjamin S. Bloom and Bertram B. Masia, *Taxonomy of Educational Objectives, Handbook II: Affective Domain*, New York: David McKay Company, Inc., 1964.

3. Kearney, *op. cit.*, p. 78

4. *Ibid*, p. 109

5. *Ibid*, p. 110

6. French, *op. cit.*, p. 131

7. John C. Flanagan, et al., *Design For a Study of American Youth*, Boston: Houghton Mifflin, 1962,

8. *Ibid*, p. 18.

9. *Ibid*, p. 130.

10. Jack C. Merwin and Ralph W. Tyler, *Nation's Schools*, November, 1966, p. 77, 79

11. John W. Gardner, "National Goals in Education," President's Commission on National Goals, *Goals for Americans*, and accompanying essays, p. 81.

12. For these and other observations by our European colleagues, see Appendixes.

13. James Bryant Conant, Shaping Educational Policy, New York: McGraw-Hill, p. 40.

14. See National Education Association Committee on Secondary School Studies, Washington, D. C., 1892.

15. National Education Association, Report of the Commission on the Reorganization of Secondary Education, Seven Cardinal Principles of Education, Washington, D. C., 1918.

16. Ibid, p. 9

17. Herbert Spencer, Education, New York: D. Appleton and Co., 1861, p. 32.

18. Seven Cardinal Principles, pp. 10-16.

19. National Education Association, Educational Policies Commission, The Purposes of Education in American Democracy, Washington, D. C., 1938.

20. National Education Association, Educational Policies Commission, The Central Purpose of American Education, Washington, D. C., 1961, p. 12.

21. Ibid, p. 5

22. Ibid, p. 6

23. National Education Association, Schools for the Sixties, New York: McGraw-Hill Book Company, 1963.

24. National Education Association, Educational Policies Commission, 1938, p. 72, 108.

25. James E. McClellan, Toward an Effective Critique of American Education, 1968, pp, 106-107, 110.

26. Walter R. Smith, Principles of Educational Sociology, Cambridge: Riverside Press, 1928, p. 58.

27. All of the cited "egocentric" and "sociocentric" definitions of education are quoted in Smith.

28. President's Commission on National Goals, op. cit., p. 6.

29. For an excellent treatment of competing philosophies of contemporary education, see George F. Kneller (ed.), Foundations of Education, New York: John Wiley and Sons, 1963, Chapters 2-4, pp. 45-92.

30. Mortimer J. Adler, quoted in Kneller, Ibid, p. 95.

31. Robert Maynard Hutchins, The Higher Learning in America, New Haven: Yale University Press, 1936, p. 66.

B. FINDINGS AND ANALYSIS

4. Establishing the Claims of Education: Requirements and Ambiguities

America was promises, in the words of Archibald MacLeish, and the promises always included reliance on education. Rush Welter who has written the best book on the idea of education is led to conclude that a "belief in popular education has been the archetypal element in our political thinking."[1] The boldness of this declaration is the more remarkable because with rare and presumably trivial exceptions -- old guard federalists, pro-slavery theorists, occasional reactionaries -- this fundamental commitment has been shared by all the schools of political and social theory that are represented in his study.

Welter explicitly rejects the thesis that ideas are "passive accomplices of personal and group interest," and he insists that American democratic thought has characteristically sought "enlightenment of the people for expansive rather than restrictive reasons."[2] We are nevertheless treated to numerous instructive instances of the line of descent that leads from interest to ideology to institutional strategy.

At the risk of doing violence to Welter's richness of scholarship and the subtlety of his argument, we may readily identify the two most prominent imputed functions of education as end points on the classic conservative-liberal continuum. The conservative impulse was strengthened by a variety of sacred and profane commitments and by the support of neutral scholarship. It sometimes relied on the routine sociological observation that a complex society requires mechanisms of social control and perhaps more often on naked motives of class interest, personal safety, and undisguised fear of the masses.

An essay in an 1838 issue of American Education combines all of the more extreme elements of the conservative stance:

> Let any man dwelling in the United States, consider this fact: that he is living in the midst of some millions of human beings, having strong bodies, strong wills, clear heads, and mighty passions; let him consider, further, that these millions suffer him to pursue his business, and sleep quietly at night, because they see it to be their interest, or feel it to be their duty to do so, but that, as soon as they cease to see their interest, or feel their duty, they may pull his house about his ears and hang him upon the nearest tree; -- and he will feel, to his heart's core, the necessity of wide-spread moral and religious education to his own safety.[3]

The conservative emphasis on education as an instrument for the preservation of order and the maintenance of hierarchy was obviously incompatible with the interests of workingmen, farmers, social reformers -- of any who spoke for the disenchanted and dispossessed. The liberal opposition cherished education as the most promising means of broadening the base of political intelligence, of facilitating the circulation of the elite, of removing the barriers that impeded equality and social justice. As Welter points out, for nearly a century after the age of Jackson, democratic theorists "treated popular education as the one sure cure for contemporary social and political evils."[4]

The most durable preoccupation of nineteenth century political thought, then, was the nature and limits of popular rule. The Spencerian thunder of William Graham Sumner could intermingle with the gentle musings of an Edward Bellamy, but each was in part a theorist of education. The solidity of this conception was shattered by the historical events and the intellectual discoveries of the recent past. The doctrine of "anarchy with a schoolmaster" is plainly no substitute for extensive governmental participation in the affairs of men; Bentham's reasonable citizen judiciously wrestling with the hedonistic calculus has been consigned to the vast subterranean depths where psychoanalysts dwell and the emergence of Orwell's Big Brother is regarded by many as a genuinely plausible prophecy.

It is not strange that academic contemporary political thought tends to regard the educability and rationality of the electorate as problematic. Indeed of the two formulations now most in vogue -- countervailing power and leadership -- one entrusts democratic decision-making to an invisible hand and the other to an aristocracy of talent. Neither dignifies mass education by assigning it the status of a major variable. Writing at the beginning of the sixties Welter concludes that "we have virtually conceded the failure of democratic political education to serve its intended purpose."[5]

Welter deplores the abandonment of the older conception of education as a viable influence in the political arena. It is doubtful that the scholarly estimates against which he reacts were ever very pervasive in the public consciousness. Certainly by mid-decade, presidents, opinion-leaders, and citizens had reaffirmed that education was vital to our national existence; the decisions reached in the marketplace of power and the councils of the elite always require ratification or at least passive consent by an informed electorate. A democratic theory that is not truly pluralistic is impoverished, for as Welter says, "we may recognize the group basis of our national life, acknowledge the irrationality and weaknesses of a democratic electorate, and accept the necessity of political leadership, yet insist at the same time that some kinds of popular education are indispensable to a democracy because some kinds of popular intelligence are necessary."[6]

The faith in education as a political force has its counterpart in an equal conviction of the positive functions of schooling in the economic sphere. The folk belief that education pays has been sustained at the macrocosmic level by a growing body of economic findings. There is considerable evidence that the rising education of the labor force, as much or more than increased investment in plant and equipment, accounts for our national prosperity.

Indeed, the crude measure "years of school completed" is a very powerful predictor of a variety of outcomes. There is a positive association between educational attainment and income, occupational status, marital stability, mental health, economic conservatism, voting behavior, commitment to libertarian values, racial tolerance, and lawful behavior.[7]

The American faith that exposure to school, any school, for a specified number of years is a meaningful index of accomplishment is demonstrated by the fact that educational attainment is the primary basis of eligibility for the overwhelming number of occupations in our society. Since it is a moot question whether there is any functional relationship between years of school completed and the actual skills required to hold a job, self-fulfilling prophecies may account for some of the benefits ascribed to the school.

It would be a mistake to conclude that education is regarded as purely an instrumental value which is prized for its contribution to social welfare or personal ambition. An excerpt from recent essays provides an an instructive illustration of the claims made in behalf of the intrinsic properties of education:

> No one is ever well enough educated about anything; that no one ever knows enough to exploit fully the possibilities of his own life or to enrich the lives of his fellows; that no one is ever well enough trained, in his tastes and sensibilities to apprehend the creative achievements of his civilization with sufficient wonder and delight; and further, that no one is ever as aware of himself and of the reaches of his capacities as he could be, nor are the capabilities of which he is aware, whether sharply or dimly, ever as fully developed as they might be.
>
> If these assertions are granted, it follows that there are no upper boundaries on the amount of education that anyone really 'needs' nor any valid general basis for establishing cut-off points at which schooling should be terminated for any segment of the population. Instead, a lifetime of continuing education becomes an indispensable prerequisite to the good life and effective citizenship.[8]

In the face of such lyricism it seems ungrateful to point out that there is a distinct paucity of evidence that education, as such, is actually responsible for many desired or unsought outcomes even those which are highly correlated with formal schooling.

This is scarcely an original observation; it has been often noted but seldom taken seriously. Almost a half-century ago the educational sociologist, Charles C. Peters, surveyed then existing educational claims and arrived at conclusions which do not require improvement. His comments merit extended quotation:

> 1. Many of the claims are entirely unsubstantiated by the facts. A frank recognition of conditions does not show that history as long taught in the elementary and secondary schools makes much contribution to good citizenship. If one will ask himself in just what way the history he has studied, or the history he is teaching, has affected or could be expected to affect one's civic reactions, he will find himself able to locate very few actual contributions. What did he learn about Lincoln or Washington or Jackson that could induce him to go to the polls when otherwise he would have stayed at home? What political event or economic principle was so discussed in his school history as to change his vote from a particular Republican to a particular Democratic congressman? The claim of contribution to citizenship seems very plausible until one presses for details; then it vanishes into thin air. . . .
> 2. Many of the claims rest upon an a priori consideration of possibilities, not upon probabilities. That the study of Greek and Latin should induce in one an appreciation of the classic foundations of modern civilization and give him direct access to a rich literature for his aesthetic enjoyment, is an abstract possibility; but no observant person would believe for a moment that it is lifely to do this with the mass of high school pupils who take the study as it is offered in the secondary schools. Nearly all the alleged transfer values are possibilities, not probabilities. The many practical applications claimed for several studies (as algebra, physics, mechanical drawing) are mostly probabilities only for the specialized few, while for the many they can scarcely be regarded as even in the realm of possibilities.
> 3. Where the claims are true they are often so vague and large as to help us scarcely at all in knowing where concretely to take hold. Studies are said to develop initiative, or self-reliance, or judgment of relations, or personal culture; but these are all omnibus ideals that illuminate the whole area but do not provide us any guiding star to lead us onward. They need to be broken up into such specific forms as will suggest to us how we should proceed in teaching them so as maximally to realize these ends.
>
> We need to replace these empirical conclusions regarding the values of the several subjects with scientific investigations that will be impersonal, systematic, observational.

> 1. Fundamentally we shall need to know, on a scientific basis, what are the values that are needed in society -- the fitnesses that individuals must have if they are to be socially efficient in maximum degree.
> 2. Next we shall need to determine scientifically what is the subject-matter that can make largest contributions toward these desired ends.
> 3. Then we shall need to know, on an experimental basis, what are the methods of handling the subject-matter that will be the most economical of time and energy.
> 4. Finally, we shall need to have some adequate measuring instruments that will enable us to tell whether or not, and how largely, we are succeeding in actually attaining these ends.[9]

A decade later, Daniel Kulp, among the most irascible of the early educational sociologists, posed the query, "Why have schools?" In a field of human experience that has been so widely organized, have the outcomes proved the value of the experiment? After an enthusiastic demolition of existing claims Kulp replies:

> We have schools because we have laws that compel us to. This situation we rationalize with reasons, some sound, some unfounded, as yet some wholly false . . . We are on a train and move because the train moves. Some of us are quite certain where the train is going; some of us are not quite so sure, especially when we wonder what some of the scenery is worth . . . .[10]

The claims of education remain dubious because most propositions about schooling are based on 1) proverbs, 2) behavior maxims, 3) collective testimony, 4) theoretical extension, 5) reports by observers, and 6) assertions by practitioners. A brief supplement on education in a Philadelphia newspaper provides illustrations of the wisdom that can be extracted from nearly all of these sources of knowledge.[11]

1. Proverbs

"As the twig is bent so is the tree inclined."
"Spare the rod and spoil the child."
"The apple doesn't fall far from the tree."

2. Behavior Maxims

Children act with each other in the same fashion in which they are treated. Fighting as a means of resolving grievances among classmates is more likely to take place in schools in which corporal punishment is consistently used to maintain order than in schools where discussion and restraint are applied.

3. Personal and Collective Testimony

He knew how to keep the course from being so dry. One of his methods was to divide the class into football teams. Each side would get four downs. The ball would be advanced if a boy was successful in his recitation. The tougher the question, the further a correct answer would advance the ball.

It was tough scrimmaging for me, but with Brecht's encouragement I passed the second year of Latin. That was enough.

I'll be jeered for saying it, but a teacher who exacts condign punishment seems to develop a deeper relationship with her pupils. Witness the frequency of statements by parents about how "Miss Anderson was my favorite teacher. She was strict and would punich me, but I liked her best. I'll never forget her!" For the worst crime a teacher can commit is to neglect her little ones. A timely flick, adroitly administered, at a pupil's ear can be the answer to the prayer, "O Lord, we who are indifferent, make us different!"

4. Theoretical Extension

When I look back and ask myself, what have I specially done for the very being of education?, I find I have fixed the highest supreme principle of instruction in the recognition of sense impression as the absolute foundation of all knowledge.[12]

5. Reports by Practitioners

"I believe in the discovery method of teaching," Mrs. Wilhelm said, "I've been using it for thirty years. I'll hint, give clues, I'll do anything but tell them. They have to discover themselves . . . that's what learning is."

We need not belabor the point that these and similar modes of comprehending the educational process cannot much enlighten us about the impact of the school. A much smaller body of discussion about education tries to establish the relationship between particular educational practices and putative outcomes of schooling by employing "scientific" methods of controlled observation. In assessing the value of such investigations, it is important to bear in mind the requisite conditions for establishing an educational claim. Educational research is beset by three inherent difficulties which renders many of its conclusions moot. These are: 1) the problem of distinguishing the effects of input variables from characteristics of the educational system; 2) the problem of distinguishing cohort effects from historical effects; 3) the problem of distinguishing selectivity effects from school effects.

Inputs, System Characteristics, and Outcomes

Even if we adopt the most economical of all models, a three-dimensional scheme consisting of 1) social or individual resources and constraints, 2) characteristics of the educational system, and 3) outcomes, a legitimate claim for the independent influence of the second of these can be made only in very few instances. For example, let us imagine that we are comparing two groups that are sometimes similar and sometimes dissimilar according to three dichotomized dimensions: native intelligence (high I. Q. - low I. Q.), conditions of instruction (big classes - small classes), and academic achievement (good grades - poor grades). Under these circumstances, 64 distinctive patterns will emerge, half permitting causal inferences and half inconclusive. As it turns out education may be responsible for the outcome in only one-fourth of the possible instances. In the following table "+" means consonance between the two groups and "-" means that they differ:

	External Resources and Constraints	Characteristics of Educational System Quality of Instruction	Outcomes
Determinant	I. Q.	Size of Class	Grades
Education	+	-	-
	-	+	+
External Resources and Constraints	-	+	-
	+	-	+
Ambiguous	-	-	+
	+	+	-
	+	+	+
	-	-	-

A characteristic of the educational system, in this case size of class, is identified as a source of an outcome in this hypothetical illustration because potentially contravening explanations "exterior" to the school system have been eliminated. Where I. Q. is held constant grades vary with conditions of instruction. Even when I. Q. differs the factor of class size is sufficiently "strong" to "overcome" the variation in "native intelligence."

The set of contingencies to which we have referred never exist as "pure types," can be expressed more elegantly in statistical form, and together do little more than affirm that a constant cannot explain a variable and vice versa. They have been introduced here because they vividly demonstrate that even when we employ an excessively simplistic model based on weak experimental design, with only three dichotomized dimensions, where each parameter is assumed to be coextensive with the sphere it represents (i. e. size of class represents in progressively more general fashion "classroom practice," "conditions of instruction" . . . educational system), and freedom from measurement error is assumed -- even within these generous limits education is a sovereign cause only in a restricted number of highly circumscribed cases.

The difficulty of controlling confounding variables that are external to the educational system may be the single greatest dilemma confronting students of the educational process and it has seldom been possible to discover what proportion of the variance in an observed outcome is attributable to the effects of formal schooling. The Coleman Report which is the first study based on a national sample which made a serious effort to confront this issue yielded disappointing findings for those who cherish a belief in the power of education.

Of the many implications of this study of school effects on achievement, one appears to be of overriding importance. This is the implication that stems from the following results taken together:

1. The great importance of family background for achievement;

2. The fact that the relation of family background to achievement does not diminish over the years of school;

3. The relatively small amount of school-to-school variation that is not accounted for by differences in family background, indicating the small independent effect of variations in school facilities, curriculum, and staff upon achievement;

4. The small amount of variance in achievement explicitly accounted for by variations in facilities and curriculum;

5. Given the fact that no school factors account for much variation in achievement, teachers' characteristics account for more than any other--taken together with the results . . . which show that teachers tend to be socially and racially similar to the students they teach;

6. The fact that the social composition of the student body is more highly related to achievement, independently of the student's own social background, than is any school factor;

7. The fact that attitudes such as a sense of control of the environment, or a belief in the responsiveness of the environment, are extremely highly related to achievement, but appear to be little influenced by variations in school characteristics.

Taking all these results together, one implication stands out above all: That schools bring little influence to bear on a child's achievement that is independent of his background and general social context; and that this very lack of an independent effect means that the inequalities imposed on children by their home, neighborhood, and peer environment are carried along to become the inequalities with which they confront adult life at the end of school. For equality of educational opportunity through the schools must imply a strong effect of schools that is independent of the child's immediate social environment, and that strong independent effect is not present in American schools.[13]

The preceding summary does not mean of course that a more humane, exciting and responsive school could not yield outcomes as yet unseen or or even imagined. For the moment, the implication of the Coleman Report, perhaps the best known study in the entire educational literature, is that faith in the claims of education will often not survive the test of rigorous methodology.

Cohort Effects and Time Effects

Education is inherently a matter of sequences and transformation and as such invites longitudinal research designs. It would obviously be instructive to trace the educational experiences of a cohort from the first day of kindergarten to the final defense of the doctoral thesis.

Large-scale researches in depth have been so infrequent partially because few investigators are able or willing to dedicate several decades of their lives to one all-consuming research interest. Limited resources, the conditions of academic advancement, and sheer boredom militate against any such exclusive devotion to a really long-range enterprise. It is presumably possible to overcome this problem by creating research organizations, which since they are themselves long-lived, are not vulnerable to the actual or intellectual mortality of scholars who are understandably eager to get more immediate results.

The second problem is more perplexing. There is no very good way to sort out the effects of history from other influences. Since World War II the world has undergone several convulsions and it is not clear how to disentangle these imperious historical events from individual or social chronology. We are reduced, therefore, to conducting Godankenexperimenten in which we imagine away all sorts of confused and complex contingencies. If, for example, we wished to measure the capacity of a social studies curriculum to modify racial attitudes in a class that entered in 1964 and graduated in 1968, how could we protect the inquiry from the "contamination" of intrusive events in the era of Lyndon Johnson, Stokely Carmichael, and the death of Martin Luther King? The study of the effects of education suffers more than most from the fact that history is, so to speak, not subject to control procedures.

Selectivity Effects and School Effects

The achievement of external validity, the extension of findings to a larger universe, is a persistent problem of all research designs. The effects of subject mortality and differential selection are especially difficult to control in educational inquiry. For example, generalizability may be weakened because schools which agree to cooperate in an experiment may differ in crucial respects from those who decline. This source of potential distortion is well understood even if it is difficult to deal with it in practice.

An equally important desideratum which is often not recognized is the necessity to assign individuals randomly especially in ex post facto experimental and quasi-experimental designs which comprise so much of contemporary research. Thus, for example, the differential earnings of elementary, high school, and college graduates is commonly attributed to length of schooling. This interpretation is sustained by comparing current income of persons who completed specified educational levels. Quite obviously there is a contaminating self-selective factor in differential educational attainment that influences orientations to worldly success and economic achievement. The presumed connection between education and income might be substantially altered if we had observed an original representative sample of five-year-old children, one-third of whom were, by random assignment, permitted to complete only elementary school, one-third to finish high school, and one-third to graduate from college. Obviously, no investigator can, or should, control this type of real-life situation for the sake of experimental purity but neither has he confirmed what has become a durable item of folk wisdom.[14]

It is extraordinarily difficult, then, to adhere strictly to the ordinary structure of proof while attempting to establish a connection between a characteristic of the educational system and an observed item of behavior. But impurities of research design are only one source of ambiguity in educational research. The findings yielded by the most impeccably conceived design may be nullified by insufficient attention to the conventional hazards of social investigation: poorly formulated definitions and concepts, unreliable methods of data collection, inadequate sampling and statistical fallacies. The summary of the literature which appears in the appendix prepared by Smith and Smith shows in how many of these respects existing research is defective. The authors' careful analysis is consistent with Benjamin Bloom's review in 1966 of the state of educational research during the preceding quarter of a century. Approximately 70,000 titles are listed in the Review of Educational Research during this period and of this immense output Bloom says "about one out of 1,000 reported studies seem to me to be crucial and significant, approximately three studies per year."[15]

All competent observers would agree that few researches are based on representative samples of adequate size, that "causal" relationships are often established either by fiat or by low magnitude correlations, and that most of the research suffers from the failure to specify intervening variables, or to explain statistical associations in the context of a comprehensive theory. Smith and Smith's observation that the most frequently imputed educational influence is "education" otherwise undefined, is perhaps the most depressing indictment of the current literature in education. Many of the key terms in discussions about schools -- e.g. intelligence, values, equity -- are employed without any clear agreement about their meaning, let alone operational referents.

There has been gratifying recent progress in dispelling the notion that intelligence is an indivisible unitary trait. The multi-dimensional character of intelligence is well exemplified by the factorial model of the "structure of the intellect" created by J. P. Guiford and his colleagues.[16] A three-way classification scheme specifies the 1) operations, "major kinds of intellectual activities or processes" including cognition, memory, divergent production, convergent production, and evaluation; 2) contents, "broad classes or types of information discriminable by the organism," including figural, symbolic, semantic, and behavior; and 3) products, "forms that information takes in the organism's processing of it" including units, classes, systems, transformations, and implications. The relationship between these parameters is still moot, and in the important cases of the cognition and retention of knowledge vis-à-vis "divergent production" or "creativity" the association is probably weak, and possibly even negatively correlated.

These subtleties are seldom reflected in the various instruments that are commonly used to measure intelligence. This is not the appropriate forum to join the great debate on testing but surely it is no longer permissible to pretend that any summary index such as the I.Q. can, for all its considerable pragmatic advantages and modest predictive powers, serve as an adequate definition of "intelligence."

The concept of values is equally complex and even less satisfactorily measured. Charles Morris has identified the value domain as encompassing "three types of preferential behavior": 1) operative values, "a way of referring to the actual direction of preferential behavior toward one kind of object rather than another"; 2) conceived values, "some object or situation is signified and liked or disliked as signified"; and 3) object values, "properties of an object considered in relation to its ability to reinforce preferential behavior directed toward it by some organisms."[17]

A number of instruments have been developed to measure one or another of these value modalities, many, if not most, based on the original Allport-Vernon scale in 1931. Important refinements have been introduced notably by Lurie, Van Dusen, Muller and Glasser, Wickert and others but all of these share in common reliance on direct questioning of subjects as the exclusive method of obtaining value preferences. Christie and Merton have commented on the dangers inherent in this assumption:

> One way of discovering values held by students, it would seem, is simply to ask each of them about their values and them tell you. But the direct question has severe limitations. For certain kinds of values the student will reply, even under conditions of insured anonymity, not in terms of values he actually holds, but in terms of those he feels he should hold. In other cases he does not know which values he holds -- he is not capable then and there of expressing them -- and he, like most of us, must be helped to express them . . . we find that students do not intend to mislead deliberately when they report what they should believe rather than what they do believe. Social and cultural atmospheres are largely created by what members of the group believe to be the appropriate thing to believe and say and do.[18]

As a result of these and similar criticisms recent researchers have relied more heavily on indirect measures of value preferences such as the "semantic differential" which permit the subject to specify ideal attributes of others. This procedure doubtless produces more "honest" answers but, like the instruments which rely on direct questioning, the responses which it yields are no more probable to predict the actual cause of future behavior. The very few efforts to establish the relationship between verbal and overt behavior by such means as comparing the subject's report of his own values and others' perception of his actions have yielded disappointingly low correlations. In view of the unsettled state of value measurement, claims pro and con about the effects of the school on changing values should be regarded with extreme skepticism.

The conceptualization of important sociological variables has fared scarcely better. The term "equal educational opportunity," for example, includes embarrassingly rich connotations. As Arnold Anderson indicates "this seemingly simple notion has several variants with rather different implications for policy" including 1) "an equal amount of education to every-one," 2) "enough education to bring everyone to a given standard," 3) "enough education to permit each person to reach his potential," and 4) "continued education so long as gains in learning per input of teaching match an agreed norm."[19]

The generally primitive state of research in the social sciences, and in education in particular, may provide some comfort to those who make claims in behalf of one or another educational practice. The typical finding about the comparative effectiveness of alternate approaches to the curriculum and classroom practices is "no significant differences," and since this result seems to contradict both common sense and experience it remains possible to continue the classic debates with all of their customary vigor. The protagonists can, however, seldom legitimately appeal to well-confirmed scientific propositions about the educational process. The existing literature fails to provide any clearcut evidence of superiority for small versus large classes, homogeneous versus heterogeneous grouping, discussion versus lectures, live versus television presentation, non-directive versus teacher-centered classes, or independent versus directed learning. The relationships between a) teacher personality and teacher effectiveness and b) student personality and student learning are inconclusive.

The poverty of reliable knowledge about the relationship between the characteristics of the teacher and educational outcomes in view of the frequently voiced contention of the need for "better" teachers is deplorable. There is no doubt that as compared to other college students prospective teachers tend to score low in standardized tests of academic aptitude and achievement. There is, however, no demonstrable evidence that SAT scores predict with any degree of fidelity how a teacher will behave in the classroom and what will be his ultimate effect on students. To establish this relationship we would need many more systematic protocols of behavior than are currently available to us and a better account of educational consequences. Among other things we would require interactive models in the description of classroom behavior. Ordinarily we ask what are the characteristics that make for a good teacher and what are the characteristics that make for receptivity to learning without ever joining these two lines of inquiry. How well, for example, does a high I. Q. Ivy League graduate function in a slum school as opposed to a teacher with more modest intellectual pretensions who is a resident of the same neighborhood. In short, we do not know very precisely how to define a "good" teacher and practically nothing about his influence on students.

The same uncertainties extend to the organization of the classroom. A case in point is the controversial problem of homogeneous ability grouping, an issue which regularly convulses assemblies of educational practitioners and theorists. The evidence which might conceivably silence one or another of the protagonists is unfortunately almost wholly inconclusive. Surveys of the literature by Ekstrom, Eash, Passow, Franseth, Kirk, and most recently the research staff of the NEA all report that the studies yield mixed and often contradictory results.[20] The passions provoked by this issue seem entirely disproportionate to any persuasive proof that ability grouping has any discernible effects on educational outcomes.

Illustrations of prevailing uncertainties about crucial aspects of educational methodology abound. The contributors to the monumental Handbook of Research on Teaching are virtually unanimous in their estimate that we do not know enough and what we know is uncertain. Kenneth B. Henderson's survey of research on teaching secondary school mathematics reflects the mood of his colleagues who conducted similar reviews on reading, social studies, composition and literature, science, foreign languages, and the visual arts. According to Henderson:

> One is tempted to admonish the reader to draw his own conclusions about the findings and conclusions of the various experimenters [who studied the much discussed tell-and-do versus heuristic methods]. But more than this can be offered even if it is negative in tone. If the consequences of being wrong are serious, one should be wary of adopting one method rather than another on the basis of the evidence available. If the consequences of being wrong are not serious, one is freer to choose. For example, if thousands of dollars are to be committed to writing textbooks which employ Method X with the possibility that the books will not sell, or if thousands of man-hours are to be committed to redesigning teacher-training curricula to eliminate Method X, the consequences of being wrong are serious. It is hard to believe that a steel mill, for example, would alter a process it has used for some time solely on the basis of evidence on a new process no greater than that available on any of the teaching methods discussed above.[21]

Henderson subsequently indicates that he is inclined to accept Dodes' conclusions about "the science of teaching mathematics."

> The teacher can not depend upon any special type of lesson, such as 'supervised study,' to guarantee success in teaching and learning . . . . There is no decisive proof that any particular method of teaching (inductive, deductive, individual, group) or any particular philosophy of teaching (teacher-dominated lesson or socialized lesson) will guarantee better results than any other method or philosophy, so far as achievement is concerned.[22]

An examination of the foreign materials included in the appendix shows that the effects of education are even less well established in Europe than in the United States. The volume of research is less and entire areas are neglected for reasons that can apparently be partly explained by differences in national character and educational systems. William Taylor reports, for example, that English schools seldom make any self-conscious effort to impart "citizenship" skills and loyalties through social studies and civics classes. Taylor speculates that the absence of direct political education may be attributed to the fact that "English political life has strongly traditional elements; no revolutionary party commands anything more than a token following, and there is a strong commitment to orderly policy and decision-making of an orderly kind, supported by the existence of a solidly based professional civil service." It does not occur to the pupils in the non-elite schools to challenge this structure; students in the elite schools are socialized into their future roles by "incorporating the individual within the 'total environment' of a residential institution" rather than through explicit political instruction. The only recent survey on political socialization in England strikes a familiar note: Abramson concludes that "on the basis of these findings . . . it is impossible to isolate the independent effects of education."[23]

Gelpi points out that there is comparatively little research on Italian education possibly because "both in the central offices and the regional branches of the Ministry of Education there is greater concern for the legal and formal aspects of education rather than the pedagogic content." The author concludes that such research as does exist has thus far been largely impressionistic rather than scientific and experimental" and that even those findings are not extensively circulated by the Ministry of Education which in the tradition of the Italian bureaucracy is "jealous of keeping for itself what information it has." Gelpi takes special note of the fact that instruction in civics, intergroup relations, and other non-cognitive areas is both rare and poorly investigated. Limited evidence does suggest that social class and familial influences on values and achievement are as decisive in Italy as in the United States. Gelpi notes that one of the few studies on intergroup relations found "that the importance of the family is far greater than that of the school in the correlation between social communication and prejudice," and that "school grades vary markedly according to social class levels."[24]

Goudsblom's summary of educational research in the Netherlands shows that much attention "is paid to the problem of whether schools prepare their students well enough for further schooling." Goudsblom attributes this preoccupation to the fact that the Dutch school system enjoys a high degree of autonomy and "is run primarily with an eye toward 'tasks' set by the system itself."[25]

Despite variations in national character and educational systems, there are nevertheless points of resemblance in the nature of research outcomes reported for all of the nations represented in this study.

1. The range of educational goals specified for each school system is far wider than the range of outcomes studied. The bulk of research findings refer to cognitive variables to the neglect of other interests. Taken collectively, they convey the impression that the primary function of each school system is to prepare the student for the next level of education.

2. Existing research has usually failed to establish the independent effects of education as distinguished from "external resources and constraints." Such variables as social class, sex, age, and "native ability" probably account for a substantial, possibly the greater, part of the observed variance in cognitive achievement, value orientation, and personality transformation.

3. There is little persuasive evidence that favors any of the warring schools of education. No array of school practices have been demonstrably more effective in achieving desired outcomes than rival procedures.

4. The shared characteristics of educational systems are probably the chief causal component in societal and individual outcomes that may be attributed to schooling. Much of the gloom about the putative effects of education may derive from the failure to observe the relationship between independent and dependent variables at the same level of theoretical generality. The disappointment with the failure of institutional characteristics such as the nature of the curriculum, the intellectual "climate," the size of classes, etc., to influence psychological outcomes such as "emotional maturity" may simply reflect theoretical naivete. After all, psychiatrists who are engaged in a direct one-to-one relationship with individual patients report a discouragingly high incidence of failure. It may well be that research in the sociology of education will be most profitable when inputs and outputs are both on the same theoretical level, that is to say, when efforts are made to establish the relationship between gross institutional measures and gross social consequences.

The concentration of research effort on the macro-level would have the felicitous effect of directing our attention to the uniform, constant, and durable properties of educational systems rather than their marginal and peripheral features. For example, American investigators have been preoccupied with detecting the consequences of "authoritarian" versus "democratic" leadership. The results of these inquiries have been disappointing partially because normative perscriptions defining classroom atmosphere severely restrict the amount of permissible variation. It is precisely these common features of classroom practice that have been ignored and therefore discounted as sources of educational outcomes.

Nevertheless one could make a plausible case for the proposition that schools as such create attitudes favorable to the "needs" of a modern industrial society. This insight is partially confirmed by the literature on education and economic development and is linked to the interest in "equality of opportunity" in all of the nations represented in this study. This concern is reflected in the development of the English comprehensive school, the French Reform Act of 1959, the recent German preoccupation with "life chances" and stirrings in Holland for a "more energetic University policy." The importance of the school in the economic sphere derives from its role as an attitude producing environment. Every teacher demands of his pupils constant adjustments and changes; some of these are small while others require discontinuous shifts to more austere skill levels. The organization by grades provides a miniature mobility model with provisions for success as well as failure. Moreover, the school necessarily requires problem-solving behavior and is typically regulated by achievement norms. It is difficult to imagine a more effective introduction to the spirit of modern economic life.

Claims of this order must, of course, be judged by the same standards of evidence as those advanced by other authors. In any event we may not be consoled by the self-evident fact that complex societies could not exist without schools and still retain their present character. We would wish reassurance of adequate "productivity" in educational institutions; i.e. a favorable ratio of imputs to outcomes. The implication of this review is that, for the most part, we do not know if schools achieve what is claimed and hoped for them and much of what we do know suggests they do not. The first temptation is to attribute this state of affairs wholly to the limitations of schools or social research or both. An alternative interpretation is that the school operates under inherent constraints which limit its influence including: 1) innate restrictions on human malleability, 2) intrinsic boundaries of formal education as part of the socialization process, and 3) temporal limits in the persistence of educational effects.

1. Human malleability

The conception of man that best sustains faith in education views him as malleable in that he has few constitutionally or socially derived characteristics that are not amenable to change. Any theory of learning, motivation, or perception which assigns primacy to intra-organismic processes that are minimally responsive to any external environment, also affirms by extension, that the school can exert limited sovereignty over human behavior.

The "instinct" and "fitness" theories of an older social biology as represented by William McDougall and Herbert Spencer would thus cast serious doubts on the potential efficacy of any educational system. But even now when traditional doctrines of biological determinism are thoroughly discredited, newer versions of the role of the genetic component in behavior suggest that in some areas education operates within narrowly circumscribed limits.

For example, nearly all sectors of articulate American opinion are committed to the idea of developing a meritocracy which features a class system that permits free social movement and offers equal rewards for equal talent. According to this model, if free universal compulsory education furnishes high quality education for all children, and intelligence is equally distributed among all strata, then intergenerational mobility should be "perfect," i.e. each class should contribute the identical proportion of sons to any given occupation. Any deviation from "perfect mobility" presumably reflects inequality of opportunity including educational opportunity.

But suppose as Bruce Eckland contends that "social classes are breeding populations," i.e. aggregates of individuals who are statistically distinct from other aggregates with respect to some gene frequencies as a result of assortative mating." This assertion is in fact supported by modest correlations -- in the order of .03 to .06 --in the measured intelligence of spouses. The significance of these considerations lies in the substantial relationship between test intelligence and various indices of socio-economic status and in the contention by some that the genetic component in intelligence accounts for perhaps as high as 70 percent of the inter-individual variance. We may anticipate that the within-class variance in intelligence will contract and the between-class variance will expand. It would thus follow that it will "become increasingly unlikely that the same proportions of children from each class have equal capacities to take advantage of their opportunities. The tendency of elites to replace themselves (intergenerationally) is somewhat insured by the nature of any system in which intelligence is a dynamic factor affecting status placement." This analysis implies that the inheritance of class membership is determined by genetic as well as social processes and that the school, even under the best of circumstances, can make a more modest contribution to the achievement of perfect mobility than is sometimes supposed.

There are, of course, standard counter arguments to offset this line of reasoning. Every responsible genetic theory now concedes that biological explanations leave much residual variability unexplained. Since we cannot know the full potential of any child until we give him the maximum chance to develop his capacities, it is empirically, and perhaps morally, questionable to proceed on the basis of a theory of limits. Nevertheless, we must be open to the possibility that even if educational research were flawless and schools superb we might be unable to confirm some educational claims for the sufficient reason that genetic factors decree that they cannot be achieved.[26]

2. Limits of formal education

Many theories in the Freudian tradition, Adler and Rank included, assume the basic personality and moral development is almost exclusively the result of family interaction and is substantially fixed by the time a child enters kindergarten. The school is thus able to effect relatively trivial alterations in crucial sectors of a student's life. However, neo-Freudians such as Horney, Fromm, Eriksen, and Sullivan, as well as academic psychologies such as behaviorism and field theory, do in varying degrees acknowledge the importance of late childhood and adult socialization and of environmental influences outside the family. The length of the "formative years" and the institutional locus of personality and values thus remains moot. Accordingly, there is no secure a priori basis for estimating the potential limits of the school's jurisdiction over the non-cognitive domain.

Recently, a number of investigators, notably Bernstein, Bloom, and Deutsch, have studied the relationship between preschool experiences and intellectual ability.[27] Collectively they have furnished impressive evidence that early childhood deprivation may seriously impede the subsequent capacity of children to develop cognitive skills. These allegations which were the scientific basis for the establishment of Headstart and other preschool programs thus assert for the intellectual realm what orthodox psychoanalysis has claimed for psychology. The situation is however somewhat different. While cognitive possibilities once lost are difficult to retrieve, it may be possible to meet this dilemma by the simple expedient of drastically lowering the school entrance age.

3. Temporal limits on the persistence of educational effects

Almost all of educational practice is based on the assumption that the effects of schooling persist beyond the student's departure from the classroom. The plain fact is that there is, as yet, very little available evidence that bears on the proposition that "education is preparation for life." For the most part we can only guess a) which effects of schooling become manifest at some other stage of schooling or the life cycle; b) which effects persist relatively intact for a lifetime; c) which effects become dissipated as a result of further maturation and experience.

Many teachers console themselves, perhaps legitimately, that schools are retroactively influential in the later lives of their students. A seemingly irrelevant item of information first acquired in school may become salient only when the child becomes a man. The phenomenon of the "late bloomer" in college lends some credence to the view that prior education may provide a base such that additional marginal increments of motivation or experience yield a desirable delayed reaction.

The available literature on higher education suggests that at least some of the effects of the total educational process are virtually complete by the end of the high school years. College students do acquire additional information and more sophisticated cognitive skills but there is little reason to believe that personality is significantly altered by collegiate experience. There is "in general change in the direction of greater liberalism and sophistication in political, social, and religious outlook" but the magnitude of the change is slight.[28]

Unfortunately, there does not now exist a single study which compares college students and their noncollege-age peers. There is, moreover, no reliable information on so basic a matter as the retention of knowledge by dropouts, high school graduates, or college alumni at given points after they leave school. It seems reasonable that many of the effects of schooling recede in time as memory fades, their relevance declines, and new experiences accumulate. All of these considerations combine to suggest that research which focuses exclusively on educational outcomes that are observable during the school years may sometimes seriously underestimate and sometimes exaggerate the impact of the school.

It seems evident that if education, like politics, is the art of the possible, we cannot ascertain the "success" or "failure" of school programs without some theoretical conception of what could have been achieved. Education has now become virtually a synonym for individual and social salvation. One of the primary tasks of behavior theory is to discover if, in fact, the school can sustain the burden.

It is, in any event, certain that as we proceed from folk wisdom to commencement oratory to insightful theoretical discourse, to the best empirical studies, our confidence in various types of educational claims progressively declines. Thus, the grand conclusion of the preceding two chapters is that even if we could construct a model of educational goals and outcomes which satisfied all criteria of formal adequacy there would be few meaningful entries that could be processed by the scheme.

References

1. Rush Welter, Popular Education and Democratic Thought in America, New York: Columbia University Press, 1962. p. 3.

2. Ibid.

3. Welter, op. cit., p. 81.

4. Welter, op. cit., p. 4.

5. Welter, op. cit., p. 6.

6. Welter, op. cit., p. 334.

7. For a handy compendium of empirical propositions on education see Bernard Berelson and Gary A. Steiner, Human Behavior, New York: Harcourt, Brace and World, Inc. pp. 436-442.

8. Melvin M. Tumin and Marvin Bressler, "Quality and Equality in American Education: A Partisan Essay," in Tumin and Bressler (eds.), Quality and Equality in Education, CRPF-148, 1966, Part VII.

9. Charles C. Peters, Foundations of Educational Sociology, New York: Harcourt, 1924. pp. 78-79.

10. Daniel H. Kulp, Educational Sociology, New York: Longmans, Green, and Co., 1933. p. 17.

11. The Sunday Bulletin, "Salute to Schools," Philadelphia, April 21, 1968.

12. John Heinrich Postellozzi, quoted in N. L. Gage (ed.), Handbook of Research on Education, Chicago: Rand McNally and Co., 1963. p. 29.

13. United States Office of Education, Equality of Educational Opportunity, Washington, D.C., 1966. p. 325.

14. See Egon G. Guba, "Common Sense About Experimental Design in Educational Research," Paper read at a Faculty Seminar, School of Education, New York University, February 25, 1963.

15. Benjamin S. Bloom, "Twenty-Five Years of Educational Research," American Education Research Journal, 3, 3, May 1966. pp. 211-221.

16. J. P. Guilford and Ralph Hoopfner, "Structure-of-Intellect Factors and Their Tests, 1966," Reports from the Psychological Laboratory University of Southern California, 36, June 1966.

17. This entire discussion is based on a fine summary article by Wilson and Nye. See William J. Wilson and F. Ivan Nye, "Some Methodological Problems in the Empirical Study of Values," Washington State University, Bulletin 672, July 1966.

18. Richard Christie and Robert Merton, "Procedures for the Sociological Study of the Value Climate of Medical Schools," Journal of Medical Education, 33, 1958, pp. 125-153.

19. C. Arnold Anderson, "Patterns and Variability in the Distribution and Diffusion of Schooling," in Anderson and Mary Jean Bowman (eds.), Education and Economic Development, Chicago: Aldine, 1965. pp. 341-342.

20. National Education Association of the United States, "Grouping in the Public Schools," Research Memo 17, July 1966.

21. Kenneth B. Henderson, "Research on Teaching Secondary School Mathematics," in Gage, op. cit. p. 1021.

22. I. A. Dodes quoted in Henderson in Gage, op. cit. p. 1025.

23. See appendix.

24. See appendix.

25. See appendix.

26. For an extended treatment of this issue see Marvin Bressler, "Sociology, Ideology, and Biology," in David C. Glass (ed.), Biology and Behavior, New York: Rockefeller University Press, 1968. pp. 178-210.

27. See Basil Bernstein, "Some Sociological Determinants of Perception," *British Journal of Sociology*, IX, June 1958, pp. 159-174; Benjamin S. Bloom, *Stability and Change in Human Characteristics*, New York: John Wiley and Sons, 1964; see Martin Deutsch and others in *The Merrill-Palmer Quarterly*, July 1964, 10, 3.

28. See Nevitt Sanford (ed.), *The American College*, New York: John Wiley and Sons, 1962. p. 106 et passim.

B. FINDINGS AND ANALYSIS

5. The Evaluation of the Effectiveness of Educational Systems: Some Problems

The need for evaluation of the effectiveness of educational systems in the United States arises from three related sources:

1. Education is a public enterprise and, as such, involves the spending of scarce and valuable public resources. By that token those in charge are legitimately responsible for an accounting of the disbursement of these public resources and of the effectiveness of that disbursement, given stated ends.

2. In a democratic society, that accountability takes on an added dimension. For it is no longer simply an internal bureaucratic matter. Rather in a society where education is "everyone's business," the adequacy of that system is also everyone's business.

3. The implied commitment of democratic educational systems to "education for change" and the volatility of social institutions in the society at large, requires that inventory shall be taken at periodic intervals to assess the extent to which there is substantial movement toward desired ends.

It was shown in Chapter I, that these three sources from which the need and the legitimacy of evaluation arise are at the same time three major sources of complexity for any one who tries to develop an adequate model for the measurement of the effectiveness of such educational systems.

Thus, the high visibility of the educational system and its governance, which underlies the public accountability of that system, also makes possible the constant confrontation, albeit partial and often inaccurate, by relevant publics, such as parents and employers, of the "outcomes" of that system, as they personally encounter those outcomes in the form of their children and their employees.

So, too, the legitimacy of and relevance of the voices and opinions of numerous concerned publics, related to the educational system in one or another degree of directness and primacy, endows their expressed concern with the outcome with legitimacy and gives them the right to argue for the relevance of criteria of adequacy other than those currently being followed by those delegated to run the schools.

Finally, the fact that education in a democracy is by that token "education for change" complicates the evaluation problem enormously by taking the fact of "shifting goals" and "shifting criteria" as a "given" of educational operations. Evaluation of stable systems, with fixed goals and fixed criteria of achievement, is difficult enough. Those difficulties are obviously multiplied many times over when the continuous "emergence" of goals is taken as a natural and desirable condition of the system.

The difficulties just cited notwithstanding, the need for and the political pressure toward the development of sound instruments for the evaluation of the effectiveness of the system are undeniable. That effectiveness may be said to be measurable by the extent to which the intentions or goals of the system are being achieved, within the limits of the stated or expected costs. Every guide to evaluation that one can find in the literature stresses the importance of getting the goals of the enterprise clearly stated. Otherwise, no evaluation is possible -- certainly no rational evaluation.

But the problems involved in discovering what are the goals of an enterprise such as the "educational system of the United States" are far more complex than they seem at first. These matters are being treated at length in another section of this report (see chapter on Goals). Suffice it to say that any one of several possible indicators of the goals could be utilized. These include the following:

1. Declared ends by chief educational officers in the system.

2. Declared ends by chief political officers of the society or community in which the educational system is located.

3. Expressions of community sentiment (or several or many of them) regarding what the schools ought to be doing and accomplishing.

4. What teachers are trained to work toward and to try to accomplish.

5. What school boast about or take pride in or compare themselves on, relative to other schools or other societies' systems.

6. What the schools measure as the mark of the achievement of the students.

7. What the schools honor, by way of reward and commendation, in the performance or behavior of the students.

8. The content of the so-called formal curriculum -- namely, what is sought to be "taught"; and here how much time, money, and skill in the school system is devoted to what subjects? This is an operational mea-

sure, in place of which one might put the school budget broken down by categores of topics to which various sums and kinds of personnel are allocated.

9. What skills, capacities and knowledge are used as bases of judgment by one stage of the school system to decide which students ought to go on to the next stage (e. g. note the college admission practices and criteria here).

10. Finally, one could look at the range of adult tasks and roles which everyone is supposed to perform and ask to what extent it is explicitly or implicitly assumed that the schools will be instrumental in preparing young people for these adult roles. This would be a bundle of intentions which would surely overlap considerably with explicitly stated intentions of school officials, community members, political leaders, etc.

Clearly, one will be seeking to determine different things about the outcomes of a school's efforts, depending on which of these various possible "goal statements" one uses as his guide.

Clearly, too, when the question is put in terms of seeking to discover the extent to which a school system is achieving its desired ends, one is implicitly committed to using as indicators one or more of the declared sets of intentions of participants in the school system. For many purposes, such a procedure would be quite legitimate and useful. One need not raise the question of whether the school should be striving after this or that goal, but only given the goals it declares for are they being achieved?

There are, however, serious complications involved in following this line of thought even if one confines his attention to only one school system. For the relevance of differing publics and their opinions as to what the schools ought to be achieving makes it extremely difficult to select arbitrarily any single set of goals affirmed by any single public or derived from any single source (among those listed earlier). Suppose, for example, that it is decided to evaluate how effectively the system is operating against the criteria of the "declared goals" of the chief educational officers in the system; and then it is discovered that there is serious disagreement in the community with regard to these goals?

The problem here is similar to, though not identical with, the comparable problem in evaluation. For the competing publics involved in any one educational system may present as wide a diversity of desired outcomes as one might find among various national systems of education even if one took as indicators the formally declared purposes of education as the measures of the "ends" of each of the systems involved. In brief, one does not eliminate the "comparative" problem by focusing alone on one educational system. At least one does not eliminate those aspects of the comparative problem which have to do with the diversity of goals, even though, in the comparative situation the diversity occurs among several systems being compared rather than among publics within one system.

This problem is present to some degree even in more monolithic systems in which the presence of differing publics is politically less relevant simply because there is a single center of "goal declarations." Such is the case in some national educational systems where national ministries operate according to formally prescribed and apparently clearcut mandates. In such systems -- traditional Italian education is one example -- one encounters a significant disparity between what the national ministry declares ought to be done and what the teachers in fact choose to do with the educational mandates and resources given to them. The teachers, as the operating agents of education, are the effective goal setters, within the framework of the very general guidelines given by the national ministry and the formally prescribed curriculum.

That "diversity" and "dissent" are always possible, even in traditional and autocratic systems of education, is evidenced in the recent waves of protest against traditional education in various European countries. These protests have been no less violent in manner or fundamental in content than those recently experienced in the United States. They indicate that the degree of dissent and demand for change is not necessarily correlated with the degree of looseness, ambiguity, and legitimated dissent that has been previously present in the system. Recent reforms of education, or the promises thereof, in such countries as England, France, Germany, and Italy, among others, reveal quite clearly that the voices of the "captive students" or clients of the educational system must now also be counted in the diversity of "goal formulators," even if they have been largely still or ignored until now.

Problems of inter-system comparison

The emphasis here on the comparative dimension of the problem of evaluation is caused by our intention to develop a model for the measurement of the effectiveness of educational systems -- without specific regard to what kind of system it happens to be. It was in line with that interest that we first developed some specifications of the boundaries, structures, functions, and strains present in any and all educational systems. It was then indicated where the problems generic to all such systems took on special meaning and dimensions. This occurs when the systems are specified and differentiated by size, number of personnel involved, number of functions allocated to the system, number of systems in interaction and type of socio-political values to which the society and the education systems were oriented.

If one seeks to develop a generic model for evaluation for any and all systems, the "comparative" problem becomes paramount. These problems have been described and considered in great detail in a number of important documents (see, for example, Comparing Nations).[1] They need not be recited in detail here. They include all the expected problems encountered in comparing across system boundaires in any one heterogeneous culture, such as 1) the comparability of the meaning of questions, 2) the comparability of indices such as achievement tests, 3) the variability in the implications for the society of varying outcomes (e. g. what does it mean to society "x" to turn out a labor force with ten percent professionals as against the society "y"? If the needs are very different the implications of this outcome, and its "desirability" will obviously differ); 4) the comparability of inferences from the same kinds of data (e. g. the underrepresentation of an ethnic group in an occupational or educational elite in one country may be due to very different reasons than those which account for a comparable underrepresentation in another country); 5) the comparability of aggregated data (e. g. a figure representing the median number of hours of homework done by school children in one country on a given subject may be approximately the same as the figure for children in another country, but the profile of the distributions may vary greatly); and 6) the "meaning" of homework may be substantially different, class by class within any one country, and across countries.

We put aside these technical problems, important as they are, in order to focus on our major problem of the moment: how does one compare nations or systems if they have avowedly different educational goals in mind? That is, assuming for the moment that all the problems of how to get reliable data relevant to comparisons have been solved, we confine out attention alone to the problem of diversity of goals. Thus, if nation "x" seeks to train an elite of technical and scientific personnel and nation "y" is mostly concerned with a vision of genuine equality of quality education for all children, of what use will it be to compare such nations and how can one make a meaningful comparison?

One possibility is to compare systems on just this feature -- namely, the difference in the goals sought by each, and the extent to which the differing sets of goals are being achieved. One might also ask what proportion of commonly desired goals, if there be any such, were being achieved in respective school systems. All these would be interesting and useful for some purposes.

What is needed, however, is a framework of comparison which not only makes all of these just recited comparisons and others possible, but also provides an inventory of all possible goal orientations against which the aspirations and achievements of any system could be measured.

Such a master list is in principle easy enough to construct. It can be done inductively by canvassing a representative sample of goal statements and including everything one finds on such a hunt. The chances that systems investigated subsequently might have still other goals is not unlikely, but it is not serious. One simply expands the list when such additional goals are encountered.

Alternatively, one can proceed somewhat more systematically, and perhaps more comprehensively, by deductively constructing a master list of possible goal strivings out of a generalized conception of the educational process and all its possible outcomes. This requires a conception of that process that is sufficiently general to comprehend most "things" recognized as a form of an educational system at the same time that it is not simply a synonymous formulation for a society in general. But this is more difficult than it sounds.

For, if an educational system is defined as that body of norms, agencies, personnel and processes which is chartered to prepare the youth of a society for effective participation as adults, and one then goes on to the possible roles here implied, one unavoidably ends with a form of analytical classification of all societal activities. This may be hidden at the outset by the level of generalization or specificity on which the roles are stated. The level may be so general that typical adult roles, normally stated in more specific terms, may not easily be located on the list. Or, alternatively, the list may be so specific that the broader terms in which general social roles are normally stated may not easily be seen as organizing principles of the detailed specifics.

These are only routine problems of classification, however. They do not materially alter the basic fact that a list of possible goal orientations of educational systems that would be both sufficiently comprehensive to include the ranges of probable empirical cases and yet be relevant to a specific conception of education (as the process of preparation for adult life) would, in fact, be a list of all major social roles normally required or potentially requirable of adults. In short, the list of all possible educational goals is the list of the structural or functional requisites of societal maintenance stated in role-playing terms.

This is quite evident when one scans the following pages labeled "Domains," "Aspects of Domains," and "Facets of Domains," which represent a list of the institutional areas and activities in which the adult roles for which students are to be prepared could be located.

This classificatory scheme is presented here as an example of what an all-embracing inventory of educational goals might look like. As with all such lists, alternative forms of classification might have been employed, either by increasing or decreasing the level of specificity or by making several categories out of activities which are now subsumed under one heading.

Group 5, for instance, includes activities connected with the absorption, understanding, and participation in science, art, philosophy, religion, play, and recreation. There is good reason to suppose that a neater classification, somewhat less extensive, might be devised here.

The same critique applies to the way in which domains have been subdivided into major aspects: A. the requirements of role or activity; B. the standards and norms guiding role playing; C. the specifications of the role, including routines and time allocations; and D. the rewards of role playing. These might have been made more or less general, and greater specificity in the subdivisions of the aspects might have been included.

But again, this is a relatively trivial critique. The important thing is that the list is sufficiently comprehensive to include any and all aspect of roles that could conceivably be relevant. If not, this is an error of classification rather than a fault in the general conception of what such a list should include.

In this scheme are included the distinctive among various modalities of the goals, that is, whether it is knowing about the role requirements, or valuing these requirements, or rehearsing these requirements (where rehearsing is summed up under the word "doing"), that is sought. This tripartite classification into knowing, valuing, and doing refers then

DOMAINS

Major Institutional Roles and Non-Role Structured Activities as Domains within Which Educational Goals Could Be Located

1. Economic: including both production or service role and consumption role.

2. Political: including the varying forms of possible participation in the political process from holding office to membership in party to voting.

3. Primary Group Membership Roles: including both those connected to kinship and reproductive functions and structures, and those found in non-kin structures and relationships (e. g. friendship, love, etc.)

4. Socialization and/or Educational Activities: including everything concerned with the activities of the person as a socializer or educator of the young on the one hand, and his own involvement in education beyond the cut-off point here used to locate educational outputs.

5. Activities Connected with the Absorption, Understanding and Various Participations in the Cultural Heritage of the Society: including science, art, philosophy, religion, play, and recreation.

6. Orientations to Self and Others: including intrapersonal, interpersonal, and intergroup attitudes, where otherwise not classifiable as specific to another role area, and intended to refer to the generalized person.

7. Physical Wellbeing: which we classify separately here, for no good reason other than that it is universally or nearly universally cited by every goal-stater and is not easily manageable under the above-listed categories.

ASPECTS OF DOMAINS
(Apply to each DOMAIN in turn)

Characteristics of Institutional Roles and Non-Role Structured Activities,
Which Constitute Subdivisions of the Roles and Activities,
to Guide Specification and Allocation of Concrete School Behaviors

A. Requirements of Role or Activity
   1. Relational (how do you have to get on or relate?)
      a. interpersonal
      b. intergroup

   2. Knowledge (what do you have to know?)

   3. Skills (what do you have to be able to do?)
      a. mental, psychological, etc.
      b. physical, manual, etc.

   4. Resources (what do you have to have?)
      a. material resources
      b. human (personnel)
      c. psychic (e.g. personality traits, etc.)

B. Standards and Norms
   1. Moral (including legal)
   2. Aesthetic
   3. Style and manners

C. Specifications of Role
   1. Routines (what do you do while performing the role?)
   2. Time-allocations
   3. Space and movement

D. Rewards of Role
   1. Property (rights and responsibility over goods and services)

   2. Power (derived from role or other rewards and not the power that is role-specific)

   3. Psychic gratifications (all those gratifications not otherwise classifiable as arising from or due to or contained in property and power)
      a. intrapersonal (e.g. self-consummatory, ego enhancement)
      b. interpersonal (prestige, esteem, honor, etc.)

FACETS OF DOMAINS*

(Apply to each DOMAIN-ASPECT in turn)

Types of Educational Goals or Aims or Outputs
Explicitly Affirmed or Desired, or Inferrable from, inter alia
Budgets of Money, Time, Personnel: Outcomes Evaluated: Honors Awarded:
Sources of Strain and Disappointment Recorded: Teacher / Characteristics
and Preparation Stressed: Curriculum Content Balance (Hours Allocated, etc.)

KNOWING					VALUING		DOING		
COGNITION	RECOGNITION	DIVERGENT	CONVERGENT	EVALUATIONAL	APPROVING	PREFERRING	REPEATING	ELABORATING	CREATING

---

* It is obvious to what extent we draw upon Guilford's thinking here. See J.P. Guilford and Ralph Hoepfner, "Structure-of-Intellect Factors and Their Tests, 1966," Reports from The Psychological Laboratory, University of Southern California, #36, June 1966.

to types of "educational outcomes" or "achievements" regarding various aspects of roles, in various domains, toward which educational effort could be directed.

This classificatory scheme provides for a) four major aspects of roles (without specifying the types of roles themselves) which could be found in any or all of b) seven major institutional or role-structured domains, with possible emphasis on c) any or all of three facets of any and all domain aspects. This creates a 4x7x3 matrix, or a total of 84 possible educational outcomes embracing all possible such outcomes. If further specified by the subdivisions of the domain-aspects, the additional multiplier of 13 (to substitute for the 4 domain-aspects indicated above) yields a total of 263 possible types of educational outcomes on which all systems could be compared.

It appears as though one is now prepared to start tabulating educational goals or outcomes in any system. But it becomes immediately apparent that one still has to choose among several possible competing or overlapping or counter-claiming lists of intended educational outcomes. For, as one seeks to make entries from various school systems on the matrix provided by the 7x3x13 elements just described, and if one is instructed to indicate for any given system which of the various boxes ought to be checked, one has to ask whose version of the goals should be used?

A cheerful and promising resolution to this problem suggests itself: ignore it -- that is, instead of worrying over alternative versions of intentions, focus alone on the outcomes of educational systems and enter these in the appropriate boxes on the master list. Against these claimed or demonstrated outcomes, one can then measure any of the sets of claimed intentions, to ask to what extent those have been fulfilled according to the best available evidence.

So conceived, the classificatory scheme becomes a device for the location of any and all educational outcomes, to be noted in such a way that, for any given system, it will be readily apparent 1) How many of all possible educational outcomes have been "evaluated," i. e. measured or claimed in some way? 2) How frequently have any particular sets of outcomes been measured? 3) Which outcomes have been relatively or totally ignored, so far as measurement is concerned? 4) Against the claims of the system, how many of these claims have any evidential basis whatsoever? Or, alternatively, against the stated intentions of the system, how many of these intentions have been examined for possible outcome?

To sum: start with a preconceived, analytically complete or almost complete model of all possible educational outcomes, stated at the appropriate levels of generalization and all sufficiently clear in their "coverage" to serve as adequate guides in coding and entering specific types of behavior on which there is research. In this matrix of preconceived possible outcomes, enter material found by searching through any and all kinds of studies of educational outcomes. These are then decoded for allocation to the proper category in the matrix and are then entered. For any given system, this provides a profile of educational goal accomplishments, claimed or actual, which can then be used for whatever comparative purposes may seem important.

Four such possible comparative purposes, among others, may be envisioned:

1. A comparison of a school's measured outcomes with its stated intentions (however the latter is decided); an estimate of the extent and type of disparity is thus provided; gaps are found; unsupportable claims are identified, etc.

2. A comparison of the disparities (types and amounts) between a school's recorded or measured outcomes and the possible totality of goals to which it might be committed and which it might have sought to measure. (This, then, is a measure of disparity between actual outcomes measured and the totality of all possible. One could also do this for the disparity between those actually measured and those claimed but not measured.)

3. A comparison of the disparity between intentions and outcomes of school "A" as against the comparable disparity for school "B."

4. A comparison of the disparity between actual outcomes of school "A" and the totality of all possible outcomes, with the comparable disparity for school "B."

These are only some of the possible analyses that could be made by the procedures suggested. The same matrix can also be used for entering intended outcomes. So, for any given school, there are 1) a list of all possible outcomes and various sets of 2) intended outcomes, 3) claimed outcomes and 4) measured outcomes. These can be compared against each other, for whatever they might illuminate; and then the disparities and/or the intentions, and/or the claims, and/or the measured outcomes of any given school can be compared with any of these four facets for any other school.

The complexity of these operations becomes evident when one seeks to identify measured or claimed outcomes of education -- where the claims are not simply bland assertions without any resemblance of trustworthy evidence. It is a twofold complexity. On the one hand, there is the enormous amount of material available by way of test scores of every conceivable kind -- starting with simple, daily quizzes given in individual classrooms and proceeding, at least in the United States, to the standard achievement tests, and then the College Boards, or PSAT's and the SAT's. What in this welter of material ought to be used?

The second complexity is found in the sparsity of genuine research investigations with regard to educational outcomes, where some effort have been made to relate some presumably relevant variables to given educational outcomes, e.g. the relations between prejudice and cognitive restriction or parochialism in cognitive functioning.

This second complexity nags. For, there is no provision at all in the matrix for propositions about the bearing of any given variable on any of the several hundred possible educational outcomes. At this point, one is moving from classification to genuine evaluation. In so doing, one makes clear the limited utility of this kind of matrix. It is one that serves principally the purposes of allowing for the location and notation, in systematic terms, of any of three possible versions of educational outcomes: 1) intended (however judged); 2) claimed (measured or not); and 3) measured. These can then be compared against each other; or they can be compared against the whole matrix of all possible outcomes; and then comparisons can be made for any given system with any other on any of the three versions of educational outcomes listed above, with special reference to types and amount of disparities.

So specified, it is now possible to assess the gaps between this model and one that would serve the purposes of genuine evaluation. First, it is clear that the ability of this model to measure the effectiveness of educational systems is limited to identifying outcomes, and the disparities between claims and outcomes, or intentions and outcomes. It can be used, if so desired, to enable systems to identify where they fall on various scales of measured achievements, or disparities between intentions and achievements, for their own internal comparison or for comparison with other systems. But the model does not permit any identification of the sources or consequences of success or failure of any aspect of the educational systems in question. Thus, schools can discover what they are not able to say about what they have done; or what they can legitimately say they have done; or at least what the students appear to be able to do. But they cannot discover from this model why or how they have achieved only this much and not more.

In short, the model falls far short of any classical model of evaluation whose primary purpose is the determination of the effectiveness of planned (or even unplanned but measured) social action toward given ends, with the further intent of rationally introducing further social action when the proper identification has been made of where such intervention ought most reasonably to occur.

As H. Hyman has put it, "Evaluation refers to the procedures of fact-finding about the results of planned social action which in turn move the spiral of planning ever upwards. It is the proper methodological accompaniment of rational action."[2]

Hyman then elaborates by noting that the prime problem in evaluation is "to provide objective, systematic and comprehensive evidence on the degree to which the program achieves its intended objectives plus the degree to which it produces other unanticipated consequences, which when recognized would also be regarded as relevant to the agency."[3]

When Hyman talks of a program achieving its intended objectives, he is referring clearly to a program of means or processes or acts of intervention of the so-called effective variables. He is not referring to an institution or agency such as a school except indirectly as the home or locus or the agency behind the program. The program is one of "means." And thus when one talks of the extent to which a program -- achieves its objectives -- one is asking the very complex question of whether this set of means produced the stated or measured outcomes. This is a very complex problem. It is the typical, or prototypical, classical problem of research design. In its "best" form, evaluation is equivalent to the full scientific experimental design.

Hyman notes the importance of a number of stages in any such design or program of evaluation, as follows:

1. conceptualizing the objectives of the program

2. conceptualizing unanticipated consequences

3. controlling for extraneous sources of change by experimental designs

4. controlling for effect of repeated testing

5. controlling for biases in the quality of response and those resulting from non-reporting

6. weighing effectiveness in the light of restricted ceilings for change

7. weighing effectiveness in terms of individual changes vs. net changes

8. weighing effectiveness by combining disparate aspects of change

9. weighing the amount of effectiveness

After these stages come those steps without which all the methodological rigor of the preceding nine stages would be without much point. For, what is next required to complete the evaluation process, according to Hyman, is understanding the findings on effectiveness by 1) describing the program; 2) describing the subjects; 3) assessing the differential effects among contrasted types of subjects; 4) assessing the contribution of different aspects of the program; and 5) conducting inquiry into the processes by which the programs produce effectiveness.

When one sees this agenda of a complete evaluation effort, it is clear that the model presented here is far from anything of the sort.

Yet, it may be argued, on behalf of this model, that 1) it is an indispensable first step, and 2) it is the only reasonable step toward evaluation of educational effectiveness that can be taken at this time in light of the available data on educational outcomes. (This second claim may be qualified by noting that it is probably eminently possible to conduct very good evaluative work with available data on certain very restricted differences in limited domains of cognitive outputs.)

The Indispensable First Step

Earlier it was seen why any effort to identify the "aims of education" for any given system must fall prey to the difficulties inherent in the fact that numerous publics make goal statements and claim right and propriety in setting goals for schools.

If one has to bypass the stage of conceptualizing the objectives of the program (as one must, given the difficulty and the absence of any conceivable non-arbitrary solution), one must then resort to conceptualizing the outcomes of the program, without of course specifying which program is meant in any particular case.

Instead, one speaks of all educational programs in their hypothetical totality. That totality is defined by the fact that everything and anything that in effect contributes to the preparation of the child for playing the total range of his adult roles is to be considered as part of the educational system. And even if no explicit attention is paid to various aspects of this preparation, it is clear that even the formal educational systems impinge on all of these possible adult roles to one degree or another.

But before any real pressure toward more complete and sounder evaluation can be developed, it is very important, given the politics of democratic systems, that awareness of the inadequacy of the current state of knowledge about educational achievements be produced. This is urged, on the assumption that rational and planned change only occurs when there has first developed dissatisfaction with the current state of affairs. It seems impossible to generate this dissatisfaction and, in turn, produce significant alterations in the behavior of educational systems until they have had some kind of rational inventory of their current achievements.

But it can confidently be said that no school systems, local, state, or national, has any very sound idea at all about the extent to which it has achieved its stated intentions -- whatever they may be -- so long as those intentions include more than certain restricted cognitive outputs. Judging by most of the available literature, every school system makes claims way beyond what it knows. There is profound ignorance among schools at all levels of organization and administration about most of their activities.

Hence, an inventory as proposed here is an indispensable first step.

The Only Step That Can Be Taken Now

The reasons just cited in defense of an inventory of outcomes as an indispensable first step also are relevant to the claim that such an inventory is the only step that can be taken now. For if rational evaluation is not possible until school systems learn they are inadequate, and if they can't reach this knowledge until their ignorance of their own accomplishments is made dramatically and politically visible, then obviously the rendering visible and dramatic of that ignorance is a necessary first step. But, just as important, the development of the model along the lines suggested is imperative because any further steps toward development of a more complete model (including intentions, means, processes, agencies, etc., not to mention cost-accounting both of intended and unanticipated consequences and processes) requires that there be available a body of data on these other aspects of educational systems that simply is not now available.

For example, one cannot evaluate the effectiveness of school programs aimed at citizenship training until there is information about the input, the processes, and the outcomes.

The problem here is that though one can conjure up a hypothetical and rational model of all possible school outcomes, by borrowing the model of the structure of all possible social behavior, one cannot do the same with regard to other aspects of educational structure and functioning.

Thus, one cannot conjure, without data, a model of the range of possible relationships between forms of teacher behavior and the emergence of creativity in children -- assuming it is known what is meant by creativity -- until the dimensions of teacher behavior as they might impinge on children's creativity have been conceptualized and a range of studies undertaken, thereby providing a range of empirical variations from which one might then conceptualize and develop a more general list of possibilities.

The problem here is essentially the same as that of conceptualizing all possible goals. In principle, it should not be more difficult to conceptualize "all possible teacher behaviors" than it is to do so for "all possible educational outcomes." Yet, it somehow seems terribly more difficult, and that is surely caused, at least in part, by the fact that "types of teacher behavior" as they impinge on types of student behavior have not been very well explored at all. Hence, there is little intellectual yeast and flour with which to bake the cake.

Moreover, such data are not likely to be sought until school systems feel compelled to know what they are doing. But schools are not likely to do so until they come to care about school processes as they impinge on educational outcomes. And, this cannot happen until they know and care about these outcomes and about changing them where they are unsatisfactory. The first steps, therefore, should help schools to come to know about themselves and thereby hopefully to come to care about their current inadequacies and possible ways of changing.

Formative vs. Summative Evaluation

Cronbach[4] has argued against the type of evaluation here proposed as a first step, on the grounds that it is in the long run uneconomic, and that the more efficient way is to have a continuous interplay between ongoing evaluation and ongoing modification of materials as they are being tried out.

However, there is only a technological but not a principled distinction between this type of summative research and the kind that Cronbach urges. When measuring the achievements of various school systems, it is admittedly difficult to keep a running inventory of ongoing school outcomes to be used for modification of existing school enterprises. At the same time, it is clear that just as one can keep "his finger on the pulse" of the economy, so too, one ought to be able in principle to keep his finger on the pulse of educational systems. Perhaps the pulse that one measures is much slower; perhaps, that is, one takes readings on the educational system at intervals of a year or more, rather than every quarter or every month. But, it should be possible, given the proper indices, to take sample soundings at frequent intervals and to develop trend lines and projections and hence warning and encouragement signals.

The difference between formative vs. summative research, then, is one of time-strategy. If four years are taken to be the time over which things are going to be done, and if a reading is taken once a year, and modifications made on the basis of that reading, this is formative research in the technical sense, though from the point of view of a shorter time period, the reading after one year is summative research, i. e., evaluation after things are completed over the course of a year.

The crucial fact for anyone planning for such evaluative research is that it will be useless to conduct a running inventory unless one also creates mechanisms for correction of identified trends. This means power and control of the kind invested, for instance, in the Federal Reserve Board, or comparable agencies, who seek to adjust the direction of the economy by measures regulating the flow of critical variables. One might conceivably look forward in the not-too-distant future to the time when such regulatory agencies, controlling sources of key educational variables, might be established, so that adjustments could be made periodically as the readings showed them to be required.

But, of course, reporting systems will be required that are at least as adequate as those governing the economy. And, criteria and indices of wellbeing of the system will have to be developed, along a number of crucial dimensions, sufficiently general to identify national outcomes, using data from a wide range of sources.

As a first approximation, one might think of the cooperation between the Bureau of Labor Statistics and other such agencies concerned with the flow of manpower, and the school systems. From these sources might come regular reporting as to the kinds of skills that are being created in the schools and what kind are needed in the labor market. There is no reason to suppose that in a very few years such cooperation could not quickly be developed, assuming it was agreed upon that the schools are a main source of manpower training of various kinds.

This idea may cause some hesitation because it seems to require the deliberate manipulation of the career plans of youth attuned to the national demand for various kinds of talents, and it implicitly converts the schools into giant centers of vocational guidance, training, and placement. But the vision here is broader than that. It includes identifying trends in book-reading, theater-going, creative writing, mental health, scientific understanding, and prejudice and discrimination in intergroup relations -- and then making adjustments in school curricula where it is felt that such interventions might conceivably be responsive to the identified shortcomings.

Problems of Conceptualizing Outcomes

We return now, albeit briefly, to a crucial feature of any evaluation effort, namely, adequate conceptualization of the outcomes with which we are concerned. Everyone warns of this urgency and everyone is correct in doing so, of course, for the simple reason that unless you know clearly beforehand and state clearly beforehand what it is you desire to achieve, there is no way of discovering whether you have achieved it later on. One cannot, in scientific decency, retrospectively select criteria of outcome and then point to the convenient results as evidence of achievement.

The problems facing investigators here are rather great, not so much because of any so-called great difficulty in conceptualizing all kinds of outcome but rather because of two other features: 1) the chronic lack of concern on the part of schools, school officials, parents *et al.*, with any outcomes of the elementary and secondary schools except those relevant to college admission; 2) the technical problems of conceptualization: namely, the translation of such concepts into operational terms susceptible of measurement.

Gagne[5] has made an obvious and useful suggestion: the goals should be stated in such a way that once they have been stated there is no intervening step of specifying content of the goal; that is already specified as the goal itself. Thus, for him, curriculum objectives become the expected capacities of students in specified domains of human activity.

This general guide seems useful enough, with two serious problems facing us: 1) how do we put these specified and expected capabilities of students on a level sufficiently general to permit easy collection of data? (a must when we are dealing with national educational systems), and 2) how do we identify the school-outcome correlative or indicator of the desired adult activity? The problem is easy for Gagne insofar as he defines these objectives in school-functioning terms, such as being able to multiply fractions, printing whole sentences, etc. But, when one asks, as one must, what school performances will be predictive or indicative of various kinds of desired adult role capabilities, the problem becomes much more complex. If a student knows how to multiply fractions, what does that tell us about his likely capabilities in various adult roles? At best, we can enter such an outcome in the cognitive domain as another accomplishment, to accompany the many there already about whose predictive or indicative value we have little or no idea.

This problem -- of the correlation of school capabilities with adult capabilities -- has hardly been recognized, since it has almost always been assumed to be "self-evident." One teaches Latin and Greek, of course, because Latin and Greek are what any educated man ought to have.[6] So one has a concept of the educated man as the desired kind of adult, and one knows the things an educated man must have before he can be so called, and hence one includes these things in his school training.

But, just as we are no longer sure of who belongs where in the class structure, so now we no longer have such firm and fixed notions as the educated man to guide us in the selection of school curriculum. This is all too evident, for instance, in the major codification of educational goals to be found in the literature. As one examines, for example, the work of Bloom, Krathwohl and associates, and sees the ranges of things schools are assumed to be doing, and asks what are the relevance of these to adult behavior, it becomes clear that we do not know their relevance.

So, two questions face the future evaluator: 1) What do you want the children to become? to be able to do? to feel? to know? to respond to? to be concerned about? 2) What do you know (and how do you know it) of the relationship between the so-called school behavior and the so-called adult behavior you desire the student to be ready for?

It becomes clear, then, that the disjunction, over time and content, between school behavior and adult behavior is a crucial problem for evaluation of school results. There is no point in measuring one or another school outcome unless it has some significance. This is the distinction that Scriven[7] has made between the so-called goals of evaluation and the purposes of evaluation. He refers here to "measuring teacher sensitivity" as a goal of evaluation, whose purpose, however, may be to improve selection procedures in dealing with teacher applicants, or in selectively culling out faculty for retraining.

In short, he is analytically partitioning the evaluative process into steps or phases, and taking the proximate goals of evaluation as against the more remote or ultimate goals, and calling the former "goals" and the latter "purposes." This is, in effect, only a distinction between short range and long range goals of evaluation. While this distinction is worth making, it must not be thought to be an important analytic discovery nor one that calls for different procedures in evaluation.

Needless to say, a vast frontier of unexplored territory in the field of the correlation of school behavior with adult behavior is there for the working or exploration. Conceivably, a good deal of this could be retrospective, in the sense that one might identify various types of school populations by their types of school behavior and then identify them in the adult world to see what correlation, if any, there is between the two kinds of behaviors. Thus, the ACT[8] staff has been investigating such important questions as the relations between school grades and various measures of adult achievement. In its now little classical study, it has shown that, on the basis of about 45 different pieces of research on this subject--

all of them of varying quality--one may say there is little or no discernible correlation between adult occupational or vocational success, as measured in a variety of ways, and the rank of the student by his grades in college. On the basis of this finding, Hoyt suggests that one ought to look into the whole question of grading to see whether it is serving other purposes, in view of the fact that it does not seem to serve the purpose of predicting differential success and failure at the ultimate point of payoff.

For all its shortcomings, this type of retrospective and secondary analysis of school outcomes as they bear on adult behavior is exemplary for its imaginativeness and for its possibly corrective force on current school practices. When one comes to understand better why school grades don't matter in adult vocational success, one will then have even more corrective force regarding school procedures.

The desire to have some greater prediction over, and some broader range of imagination regarding, education than has traditionally been indicated in typical ETS school measurements must surely have been among the impulses that have led to the development of the National Assessment Program. Surely the same forces earlier impelled such organizations as the Educational Policies Committee to drive toward a sounder evaluation of actual outcomes than has been possible heretofore.

The variable procedures of these two efforts -- the National Assessment Program and the work of the Educational Policies Commission -- are illuminating insofar as they indicate the predictable kind of data available for retrospective analysis of the effectiveness of American education.

The Educational Policies Commission sought first to set boundaries on the apparently unbounded and imperial claims of each of the school disciplines. They discovered that every discipline made sweeping claims for the generalized as well as the specific benefits for student development that would accrue from the study of their subject matters. The Commission argued, correctly, that with such imperial claims, no testing could be made of outcomes of any sensible sort, and indeed no corrective evaluation of success or failure could possibly be achieved.

Their first step, then was to try to persuade curricular devotees to modify and reduce their ambitious claims. One can only sympathize with this effort at the same time that one has to be very cautious about the possible over-reduction of claims. In their efforts to eliminate overlapping in claims of results, the Commission was unavoidably driven by the logic of the matter to eliminate anything beyond the rather narrowly described specialized skills for the various curricula. In short, in search of operational specifications they have risked the depletion of much of the meaningful content of the claims of various disciplines. As Cronbach put it in another context, the preoccupation with reliability can drain away evolving test-content reliability. Or, as Tyler has implied, the concern for reliability coefficients of a high order may be destructive since the typical reliability coefficients refer to individual scores and not to the homogeneity of a given level of behavior. Or, as Robert Stake has put it, the units of output treated in bad evaluation studies are often more apparently important and significant than those evaluated well.

The National Assessment effort has approached evaluation in the more traditional way. Yet there are problems here too. For, while there is a good deal of specification of a range of behaviors desired from children in the so-called hard subjects of math and science and even of English, once one goes beyond the limited skills involved here, and goes into areas such as art, citizenship, and intergroup relations, the quality of the measures and indicators drops sharply.

It is evident, then, that if a program for the measurement of school outcomes over a range of possible school goals is to be effective, there has to be equal concern with the range of possible outcomes and equal concern, then, with conceptualizing and operationalizing them for possible measurement.

Norman Kurland[9] has succinctly analyzed both some of the major reasons why an adequate system of educational evaluation has not yet been developed and some of the specifications for measures of educationalperformance that must be followed in any sound future efforts at evaluation.

Commenting on the reasons for the failure to develop evaluation to date, he argues that five main obstacles have been: 1) the difficulty of the task; 2) the belief that the important outcomes of education cannot be measured; 3) the inadequate knowledge base; 4) the lack of data and the capacity for handling voluminous data; and 5) the lack of interest in and support for education.

Regarding the specifications for measures of educational performance in any future evaluation, Kurland lists the following characteristics that any good measure should satisfy:

1. Differentiate among things that are different. Initially measures may be able to measure only large differences. Even this would be useful. As they are developed, however, they should indicate with precision smaller and smaller differences until they reveal differences that are not otherwise detected or not now detected until much time has passed.

2. Be unresponsive to changes that are not related to performance. Not all changes affect performance; a good measure should take this into account.

3. Accurately reflect changes in performance within a system and be so calibrated that similar degrees of change among systems result in similar changes in the measures. This characteristic is essential if there is to be meaningful comparison from one system to another.

4. Be capable of analysis into component elements so that it is possible to determine what specific factors are producing changes in the measures. This is essential if they are to be useful as guides to decision and action.

5. Be as nearly contemporaneous as possible with the events it reflects. Measures that reflect changes months or years after they occur are not much use for decision-making. To be most useful, the timing of reporting must be proportionate to the effective reaction time. To learn that the temperature in a boiler has passed the danger point after it is too late to turn down the heat is not helpful; nor is it much help to learn that an entire class has not been taught fractions after it has passed on to the next grade.

6. Be readily understood by those who are to use it and as free as possible from factors that may lead to misinterpretation. If the measures are to be used by busy school administrators, school board members, and others charged with responsibility for educational decisions, the measures can be presented to them in such a way that their meaning can be readily understood. This will require considerable effort on the part of those who develop the measures. They will have to reduce complex data to a few basic terms and provide an explanation that is free of technical language and directs attention to the essential factors revealed by the measures.

7. Provide short-run data on the system during the development process. Most good measures will take quite some time to develop. Development should be so designed that some information about the system can be drawn off on an interim basis. This condition is sought both to provide a check on the value of the measures and to satisfy the current needs of those who must provide the resources for development.

8. Be proportionate in complexity and cost to the thing being measured. There is no point in spending thousands of dollars to develop a measure for the effectiveness of a program whose own budget is of the same dimensions. A measurement that makes it possible to divide a class into ten groups is of little use if all the teacher can manage effectively is three groups.

Kurland then proceeds to indicate some of the kinds of data that will be needed if measures of performance that meet the specifications listed above are to be developed.

"One additional difficulty with past evaluation efforts was that they tended to focus almost entirely on what can be called, using economic terminology, "output"-test scores, number of scholarship winners, number of students who go on to college, etc. Seldom was such output data related to "input." Thus the fact that 90 percent of the students in a given high school go on to college may be largely the result of the kind of community the school is in and not at all be influenced by anything it did. Similarly if our only measure of the effectiveness of a new practice is final test scores we know little about whether the change was due to the innovation or to some other factor influencing the input.

Thus both input and output data must be carefully collected and related to one another. If there are outcomes revealed in the output that could not be anticipated from the input data, the difference can reasonably be attributed to the system under analysis. To determine just what in the systems made the difference further data on the system or "process" measures are required.

One thing is very clear from our experience to date: neither input nor output can be meaningfully discussed as though they were single, simple factors. There is much more that determines outcome than individual intelligence (which itself is a complex of factors); and much more to outcome than grades or getting into college. A successful measure will be the composite of input and output factors. What seems to be required at this juncture is to identify as many factors as seem relevant, collect as much data as can be obtained, combine the data in various ways with various weights given to the factors, and then see which produce indicators that correlate highly with subjective judgments and other measures of performance.

But the selection of output data must be related to the goals specified for the system. Whether or not, for example, reading test scores should be considered as output data would depend on whether reading achievement were a goal of the school.

Similarly the bases for selecting input and process data would be hypotheses about what factors may be relevant to the output.

This discussion suggests why an evaluation system in education must be based upon some agreement on the goals in terms of which performance is to be assessed and on sufficient knowledge of the processes of education so that meaningful hypotheses about relevant input and process factors can be formulated."

It is evident from statements such as Kurland's that the effort at sound evaluation of educational systems is not lacking for guidelines. One knows what has to be done and how it has to be done and why. What is lacking, however, is the widespread awareness that such evaluation is required. Once that awareness is developed, it may then be possible to begin to muster the required resources, including at the outset the determination to do what is required for the effective conduct of education. The creation of that awareness and determination are thus the crucial next steps.

Doubts and Hesitations

In previous pages it has been strongly urged that it is crucially important to take an inventory of educational outcomes, even if only partial, as a first step in the development of a model for the measurement of the effectiveness of educational systems. Some of the elements in one possible model of such an inventory has been indicated. The objective of such an inventory needs to be underscored. It is the indispensable first step in revealing the importance of developing a fuller evaluation scheme for measuring the effectiveness of educational systems. For, in the process of trying to take an inventory of outcomes, several things are likely to become clearer to educational systems everywhere:

1. how difficult it is to decide which outcomes ought to be measured;

2. how difficult it is to measure outcomes of any kind;

3. how limited a number of outcomes are now being measured, and how relatively badly they are being measured;

4. how partial and incomplete is an evaluation of the effectiveness of education that contents itself with the measurement of outcomes at the time of school completion;

5. how little is known about a) the factors that contribute to the outcomes and b) the factors in later life to which the school outcomes contribute.

There can be little argument with the assertion that the periodic taking of inventory of outcomes will generate fuller perception of existing inadequacies. There can be serious argument, however, as to whether school systems ought to care about these other bodies of information regarding their operations. Why, after all, should one care how the results are being achieved and what they may imply, so long as one knows the results themselves? This hard-headed, albeit narrow, pragmatism is appealing on the face of things. But a moment's further reflection reveals the shortcomings of this point of view.

At stake here, of course, is the essential question of what one means by school effectiveness. Reference must therefore be had to the conception of the educational system: what is the educational system and what are its intentions? What functions is it designed or intended to serve?

Even the apparently simplest of answers -- that the schools are designed to train children for successful adult life in their society -- reveals that the implications of school outcomes for later adult behavior are part and parcel of the "outcome" of education about which one wants information.

If, too, one wishes to know something about that adult behavior other than limited versions of occupational placement, it becomes crucial to develop ways of measuring the traditionally unmeasured dimensions of educational outcome. These include results concerned with the affective domain and with the development of sensibilities and responsiveness to the products of the culture and with citizenship and peaceful coexistence with diversity. It is the obvious relevance of these to adult life that makes it so important that they too be included in the evaluation scheme.

If, too, one is concerned with costs, then the analysis of the means or processes by which schools are achieving the school-measured outcomes, and their relative costs compared to alternative ways in which these or better outcomes might be achieved, becomes immediately salient.

Finally, if one wishes to compare the effectiveness of different schools that invest differing amounts of resources and start with different student populations, one must turn his attention to the interconnections of input, process, and output.

In short, any rational concern with achieving the announced purposes of education and even a modicum of concern for the rationality of the system lead one to an awareness of the importance of uncovering the kinds of ignorance in which educational systems currently flounder. It is toward that end and the attendant development of concern for securing the knowledge that might banish that ignorance, and the more effective utilization of that knowledge that the need for an inventory of outcomes is advanced.

The Purposes of a Full Inventory

What is being sought in the long run is a kind of an information system that would enable education in this society or any society to know where it stands at any given moment and what it must do to move more effectively toward its stated ends, whatever they may be. Such systems of information are especially needed in a democratic society, with a commitment to training children for democratic life, and with a recognized responsibility to the resource-providing republic.

In such a democratic society, the purposes of such an information system might be stated as follows:[10]

1. to help every student assess his own progress so that he can become increasingly mature in understanding himself, his educational needs, and his future possibilities;

2. to help teachers and administrators determine how effective their various programs are, so that they can take the detailed steps necessary to the strengthening and revision of these programs;

3. to make it possible for the educational authority to identify inequalities in educational opportunities in various parts of their system and to utilize their funds and resources accordingly to rectify the inequalities;

4. to provide research organizations or departments of the educational authority with the data needed to generate and test hypotheses regarding educational process, input, and outcome.

5. to provide school systems with incentives to experiment, in controlled and evaluated ways, with new and promising educational programs, materials, devices, and organizational arrangements;

6. to provide the interested and resource-providing publics with periodic and readily interpretable information concerning the progress of education in their communities.

These specifications of what an adequate information system might contain reveal both how inadequate to that purpose is the measurement only of outcomes (as we have suggested as a first step) and how vital for educational and social purposes is that broader reach of information.

Out of an awareness of these two important facts about current educational evaluation we have here advanced the argument that the determination of the effective amounts of ignorance about their own operations by school systems is an indispensable ingredient in the opening up of the question of evaluation and information-systems that would be required to serve basic educational purposes.

It can, of course, be argued that the uncovering of ignorance is no more likely to impel systems to take measures to increase their knowledge than have previous "assessments" and "evaluations" been able to achieve. That may very well prove to be the case. For there are numerous structural features in American education that make such an escape from self-knowledge eminently possible, not the least of which is the freedom from critical scrutiny by outsiders of the operations of any local school system. So long as school authorities and agents continue to evaluate themselves, there is not likely to be a very significant drive toward refurbishment, enrichment, and modification of the school system. It is therefore clear that if the taking of inventory of outcomes is to serve the purpose of increasing the incentive toward more information and the willingness toward revision, that evaluation must come from the outside, hopefully from authoritative sources, using criteria agreed to beforehand, and relatively free of any serious taint of partiality or bias.

The Limitations of the Proposed Inventory: Some Field Trials

This returns the discussion once again to the adequacy and suitability of the inventory of outcomes proposed here. For if the instrument designed to reveal ignorance and uncertainty is itself not very usable, or usable only in very limited contexts, and only at great difficulty, the intention of such an inventory may be disserved or destroyed at the very outset.

For these reasons, we have submitted our model of school outcomes to the widest possible kind of criticism and to tests of usefulness. These include a very critical examination, conducted by two members of our staff, who were charged with the task of applying our scheme of recording to educational outcomes in American society. Their lengthy report is appended here (see Smith and Smith).

We have also attempted to treat one statewide school reporting system, in terms of this scheme, attempting to identify the problems involved in the collection and aggregation of official data from the numerous constituencies who report to the state authority (see Krauss and Waldron).

We have also had a number of educational-research specialists examine our scheme against the criteria of adequacy and usefulness, with special attention to the internal features of the scheme. And finally, we have tested the usefulness of the scheme on a cross-national basis by asking educational specialists in five different countries -- England, France, Germany, Holland, and Italy -- to attempt to apply our scheme to their countries. Once this task had been accomplished, we then held a lengthy conference with our foreign colleagues to determine in detail what were the difficulties they had encountered.

In the "Smith and Smith" and "Kraus and Waldron" documents appended here one will find lengthy recitations of the details of the difficulties encountered in the application of our scheme to research findings on education and to official reports. They need not be summarized in any detail here. But it is evident from a reading of those reports that our scheme is inordinately clumsy and difficult to apply and use with any precision or reliability. This is on the assumption that suitable and relevant data are available.

Even more problematic, however, was the fact that much of the data encountered does not lend itself easily to being recorded in accordance with our categories and coding guides.

There is little chance that one could impose upon the educational research community the requirement that their reporting shall satisfy the needs of this or any other reporting system. Such changes as might make research reporting more useful to any given system of notation would involve the kind of slow, evolutionary, and impersonal shaping of the field of research that is characteristic of the scientific research community. Anyone who has tried to develop a "propositional inventory" of the state of knowledge or "state of the art" in any given sub-specialty in any of the social sciences concerned with education will readily recognize the problem presented by the great diversity of interests, methods, and styles characteristic of the total body of research literature.

In the case of the state reporting systems, there is so little concern with many of the dimensions of possible school outcome that only the smallest number of our possible outcomes are involved. Moreover, it is quite clear that a more adequate system of state reporting would be enormously difficult and complicated and perhaps politically impossible to institute, given the degree of local autonomy in American educational structure.

The criticism we solicited from psychometric experts regarding the internal cohesion and amplitude of the model of educational outcomes were no more encouraging. Serious doubts were raised regarding the comprehensibility of the scheme; the usefulness (i. e. "is it subjectively compelling for day-to-day operational use?"); the amplitude of the scheme (i. e. in some regard it was found much too general and in others much too detailed); and the possibility that tests for given outcomes would be developed.

The first query or objection asked whether the scheme was readily self-explanatory and usable. As the critics saw the matter, our scheme was deficient in this regard. There was uncertainty as to what outcomes would be located in which cells. Some of the cells seemed more "catch all" than others. There was considerable overlap that could not easily be resolved. Some of the cells seemed more meaningful at certain levels of education than at others (e. g. at professional school more than grade school).

The second query, regarding the subjectively compelling character of the scheme, such that it would commend itself for day-to-day use, raised the following questions, among others: would curriculum developers find the scheme useful? Could teachers be readily interested in taking stock of their outcomes in this way? Could the whole idea be readily put into school operations without serious disruption or addition to the already overloaded burdens of school agents?

The third query, regarding amplitude of the scheme, raised serious doubts regarding the number of cells, indicating that probably too many cells are present for efficient communication to school personnel or for a measurement "package" of reasonable size.

Finally, the fourth set of questions concerned the development of adequate tests for each of the cells. As the critics indicated, the probability that such tests could and would be developed depended on how "appealing" the cells themselves were to the applied research community.

If the scheme seemed reasonable, by criteria intrinsic to the field of educational research, there would probably develop a considerable interest in creating the necessary instrumentation. But the questions raised in the first three cases threw doubt on the likelihood that the scheme would have this kind of appeal to the test-development community.

These and similar objections must be considered when one comes to assess the likelihood that the application of this scheme to this or any other educational community might have the effects intended.

REFERENCES

1. Richard L. Merritt and Stein Rokkan (eds.) Comparing Nations: The Use of Quantitative Data in Cross-National Research, New Haven and London: Yale University Press, 1966.
2. Herbert H. Hyman, et al., Applications of Methods of Evaluation, Berkeley, California: University of California Press, 1962, p. 3.
3. Ibid, pp. 5-6
4. Cronbach's views are briefly summarized by Michael Scriven in "The Methodology of Evaluation" in Perspectives of Curriculum Evaluation, Ralph A. Tyler, Robert M. Gagné and Michael Scriven (eds.) AERA Monograph Series on Curriculum Evaluation, No. 1, Chicago: Rand McNally & Co., 1967, pp. 39-43.
5. See Robert M. Gagné, "Curriculum Research and the Promotion of Learning" in Ralph A. Tyler, Robert M. Gagné, and Michael Scriven, op. cit.
6. Alternatively, instruction in these topics or subjects is justified by some on the grounds that they are basic skills or serve as the basis of other learning, by virtue of the mechanisms of transfer of training. But this claim is, of course, insupportable.
7. See Michael Scriven, "The Methodology of Evaluation" in Ralph A. Tyler, Robert M. Gagné, and Michael Scriven, op. cit. pp. 40-43.
8. See Donald P. Hoyt, The Relationship Between College Grades and Adult Achievement. A Review of the Literature. American College Testing Program Research Reports, No. 7, September 1965.
9. Norman Kurland, "Developing Indicators of Educational Performance," a lecture presented at the 31st Educational Conference sponsored by the Educational Records Bureau, October 1966.
10. Private memorandum from Henry Dyer, Educational Testing Service, Princeton, New Jersey, 1969.

B. FINDINGS AND ANALYSIS

6. Cross-National Comparisons: Some Problems and Solutions

The general value of any model of a social enterprise is proportional to its applicability to a range of varying cultural situations. It is important, therefore, that a model for the measurement of the effectiveness of educational systems should be adaptable, without significant change, to the systems of various countries. This requirement formed the focus of several months of effort to determine the extent to which the scheme for recording educational outcomes which seemed moderately useful on American materials might also serve equally well when used to record educational outcomes in countries so diverse as England, France, Germany, Italy, and Holland.

Diverse though these countries are in some regards, they are all part of what is usually called Western European culture. Relevant to the educational system, this common cultural membership implies a number of things including 1) a commitment to education of children for participation in modern, industrial life; 2) some common, though varying, commitment as well to the idea of a democratic society in which young men and women must be fit to live and participate; 3) a corrollary ideology of educational recruitment which stresses the fitness of the student for educational opportunities rather than his "inherited" or "ascribed" right to those opportunities; 4) a strong central emphasis upon cognitive skills as the core of educational intention; and 5) some degree of flexibility with regard to curriculum so that as the changing needs of the society are defined the curriculum is altered accordingly.

These and other elements in the educational systems of Western European countries are present in very varying degrees, though they often represent more of an ideological stand than an actual set of practices. Even as such, however, they define what these systems officially "strive" toward and claim to value. Hence failures or deficiencies in these regards are bases for the arousal of educational conscience. Insofar as the systems are to that degree alike, it is possible to treat them as examples of a common educational orientation, and as dissimilar from the educational systems of other cultural areas in which these themes are significantly less present.

On the basis of these assumptions, it was decided to choose, as "comparative" cases, a sample of such Western European countries in which the suitability of the scheme of notation of educational outcomes might be tested. This choice involved a deliberate decision to limit the range of possible differences among the systems to be compared, so that at least some of the expectable variability might be kept under control. It was felt there would be sufficient differences even among such systems to afford some substantial evidence about cross-cultural suitability of the scheme of notation. It can be said confidently at this point that our expectations in these regards were not disappointed at all.

It will be of some interest to those who venture comparable comparative efforts in the future to be aware of the difficulties that are likely to be encountered in such enterprises. Perhaps the most general guideline that can be suggested is that all of the difficulties normally encountered in research within one country are encountered in comparative research in several countries, but these are compounded several times over by the differences in language, styles of research, and availability of comparative materials. Thus, for example, the problems of developing common understandings with colleagues to assure reliable procedures in coding and contest analysis are difficult in any research but in cross-national collaboration linguistic differences severely aggravate the problems. This problem is compounded still other times over, both domestically and abroad, when the guides to coding research materials into a matrix of common categories are ambiguous and indecisive. This was the experience in this research, and it probably was a function as much of the materials encountered as of the coding scheme itself. For, neither here nor abroad is the research on educational outcomes organized in terms of any preconceived logical organization of outcomes such as we here attempted to impose upon the available data. Several different sets of outcomes are thus often found discussed in the same research; similar outcomes are found hidden under different names; the outcomes investigated are sometimes so different from those provided for, that a good deal of "coding imagination" is often required to decide on the proper categories in which to enter the outcome; and, most distressing of all, most of the research focuses on restricted areas of cognitive output.

In meeting the problems of intercoder reliability, it was discovered as might have been expected that there is no effective substitute for prolonged and intense face-to-face interaction. Our efforts to develop effective collaboration through the mails and through one brief personal interchange between one of our principal investigators and our foreign collaborators were only partially effective. It was only after a ten-day face-to-face conference that certain sets of common understandings began to emerge and that certain significant sets of differences in approach and research style were clarified and adjustments made.

The sequence of events was as follows:

1. Our first step was to develop a scheme of notation on American materials and to test its worth by intensive analysis of the applicability of the

scheme to various kinds of American educational situations. These "applications" did not by any means eliminate all the problems. Rather, a more extensive understanding was developed about the dimensions of these problems and some possible ways in which to resolve these. (The range of these difficulties is recited in detail in the appended papers by Smith and Smith and by Kraus and Waldron.) That is to say, we learned some techniques for dealing with ambiguity and uncertainty in the research data, and managing, or at least "putting up with," reports on educational outcomes in whose reliability we could not have the fullest confidence.

Our accommodations to these shortcomings in the data would properly be viewed as impermissible from any purist point of view. But it was felt that at this point in the attempt to develop a model for evaluation, crudeness, roughness, and inequality in the available data had to be tolerated. Our main technique for management of these pervasive ambiguities in the research data involved continuous consultation among various coding judges until an agreeable solution was achieved. Obviously, this is an expensive and time-consuming measure. But the investment was felt to be worthwhile, on the assumption that once enough instances had been dealt with in this way, a reasonably standard set of coding guides would emerge, and the necessity for continuing consultation would decrease.

In fact, however, as time went on the necessity for consultation kept up. Though some "typical" kinds of research findings had been codified, new difficulties kept cropping up that required the same amount of subjective judgment in each case as had been required earlier. Rules that seemed applicable in earlier cases now seemed not to apply in newer cases. Overall, some progress at standardization was achieved, but not nearly the amount that one expects from the investment of time that was made. And, though the "reasoning" behind certain coding decisions was put in written form for later reminder, this reasoning did not seem decisively persuasive in later cases, since alternative reasoning often seemed equally appealing.

One example will serve to illustrate some of the persisting difficulties encountered. Suppose we encounter a research finding regarding the differential conditions under which school children learn to perform certain mathematical operations more quickly than otherwise. Focusing alone on the outcome -- "learning how to perform mathematical operations" -- one had to decide whether this was "knowing" mathematics or "doing" mathematics. For the classificatory scheme calls for a distinction between "knowing" and "doing." But if a student knows how to "do," clearly both "knowing" and "doing" are relevant. So, too, if a student learns how to appreciate a poem more deeply, is this more knowing or doing or valuing? The term "appreciate" suggests valuing. The term "learning" suggests "knowing." The active verbal implication of "appreciate" implies doing. Nor will it solve any problem to allocate this outcome to all three categories, since that in effect is to destroy the value of the distinctions.

This illustration is cited simply to reveal the kinds of difficulties encountered even in continuing and close interaction among a number of coding judges in the local research team. From this it can be seen what problems one could expect when one asked a number of foreign collaborators to follow the same scheme, and to try to do so without very much sharing of problems and experiences beforehand. To be sure, our collaborators were provided with such information as we had available regarding kinds of problems met and kinds of solutions attempted. But it is highly unlikely that sufficient similarity in coding procedures was achieved by collaborators in the various cooperating countries to ensure any satisfactory degree of intercoder reliability.

It is strongly suggested to any investigators who make similar efforts in the future that much more time and effort be devoted to close personal interaction with foreign collaborators to the resolution of the range of possible problems that are sure to arise. (All these hesitations and doubts about intercoder reliability that have just been indicated should be kept in mind when certain of the substantive outcomes are later reported.)

2. While working out our own problems on the domestic materials, it became necessary to begin recruiting foreign collaborators. We wanted educational specialists in a number of countries and we had a range of "acceptable" countries chosen, so that if assistance could not be secured from one, we had acceptable substitutes. At this point we encountered a problem that, while varying from one discipline to another, represented a sizable obstacle indeed in this case. It was the fact that the cooperating specialists had to be, if at all possible, scholars who were trained both in sociology and had done research in education. For the scheme of notation we had adopted was essentially sociological in character, with its emphasis on adult roles for which students were being trained, and its further emphasis was on sociologically conceived aspects of these roles. At the same time, the individuals who were to work with us had to be knowledgeable about the educational research in their own countries, and to have the kind of orientation to empirical research in this field that would make it possible for them to be at ease with this kind of venture. This involved, then, securing individuals who were both research sociologists, familiar with surveys, coding and the like, on the one hand, and educationists, familiar with the educational research literature of their countries.

In view of the newness of the sociology of education as an enterprise, not only abroad but in the United States as well, it was predictable that the choice of collaborators would require stressing one set of qualifications over another. It was agreed beforehand therefore that where a choice had to be made between an individual who was a sociologist but not quite thoroughly versed in the educational research materials, we would prefer this kind of research person to one who might be fully cognizant of the range of education research in his country, but whose major professional interests were only peripherally in sociology.

Extensive inquiries in the community of internationally known and knowledgeable scholars enabled us to develop a panel of names of possible collaborators in each of six different countries. It then became necessary to make initial contact, first by telephone and then by mail, with these possible collaborators to determine availability and interest. Anyone who has tried to do this and do it reasonably quickly is aware of the difficulties involved in exploring complex problems of availability and interest by international telephone. Expectably, then, the process of securing collaborators was slowed up considerably when several exchanges of mail were required to determine mutual agreeability between the domestic staff and the foreign collaborators.

It is not out of place to quote certain culturally-specific difficulties we encountered that may not recur in the same form in the future, but are not unlikely to occur in some similar form either. The fact of the matter is that we were trying to construct an international team just at the time that student protest movements were growing throughout Europe and, as it turned out, effective collaboration with both the French and German scholars was reduced to the barest minimum as a result of the various disruptions of the societies in question both by students and organized labor. Our German collaborators "disappeared" for some months; our French collaborators were most difficult to contact. We were able, however, to maintain reasonably effective liaison with our collaborators in England, Holland, and Italy.

When the best available panel of collaborators had been secured -- albeit after several months of effort -- the process of fuller exchange regarding the substance of the research effort was begun. It was at this point that the breakdown in comparative sociology began. It is difficult for us to assess the situation adequately, but it is quite clear, from the final results of the collaboration, and as can be seen from the appended reports from each of the countries, that we simply did not establish effective communication with three of our five collaborators.

The report submitted by the French team bears little resemblance to what was requested, however valuable as a document on developments in French education it may otherwise be. The Dutch documents are highly competent efforts at an overview of recurrent problems in Dutch education with very special effort to the problems of classifying educational outcomes, but concentrated primarily on the preliminary stages of the conceptualization of the scheme of notation, rather than upon the notation of research itself. The German documents are also most useful and competent materials regarding German education, but again there is little by way of detailed comparative notation that might reveal anything significant about the applicability of the notation scheme across cultural lines.

By contrast, the English collaborators produced a comparative research effort that matches point for point the effort made domestically and, in that regard, is extremely useful and helpful. Moreover, its high quality is significantly enhanced by its obvious generous command over the relevant research literature and allied efforts in British educational materials.

Working with a considerably smaller body of materials, most of which could hardly be used for comparative purposes except by generous stretches of the imagination, our Italian collaborators nevertheless managed to produce a body of comparative materials that are excellent for our purposes and probably represent as apposite an achievement as could have been wrought by any Italian scholar, given the rather forbidding limits of existing Italian research on education.

These remarks are intended, it should be clearly understood, to clarify the variegated outcomes which are appended herewith. Without this explanation it would be impossible for the reader to understand how it comes about that out of five collaborative contributions, a third devotes its time, with high intelligence, to a critique of the scheme of notation on theoretical grounds, and two others are primarily useful summaries of various aspects of education in their countries.

It is tempting to assign the responsibility for the breakdown in communication to our own shortcomings in transmitting our purposes to our collaborators. This would both "explain" what had happened and at the same time lift any pall that had unintentionally been cast upon the possibilities of international comparative collaboration in the future. Unquestionably, too, a substantial part of the responsibility must be laid at our door, insofar as we operated with the bland and sanguine notion that minimum contact and instruction would suffice to secure effective collaboration.

But there are several other factors at work that probably ought to be taken into account as contributions to the breakdown in collaboration. We refer, first, to the fact that there are strong individual styles in research that are exhibited by scholars without particular regard to nationality. The diversity of scholarly styles seems as great within any one research community as it is across several such communities.

Secondly, there are identifiable national styles and impulses in research that must be taken into account. The empirical tradition, as it has developed in the United States and England, is significantly less self-commending to a number of European scholars, at least in the special style of such research that has come to be typical of the American community.

There is third the fact that when one tries to "impose" a classificatory scheme upon the research materials from countries with diverse research traditions and diverse strengths in their research communities, one unavoidably encounters differential appeal and suitability of the scheme. In short, the same problems of ambiguity and misfitting of the research done in America by a widely scattered and far-from-homogeneous-minded research community were encountered in some of our collaborating European countries.

The combination, then, of different personal and national styles of scholars and different research traditions in the countries makes the kind of collaboration we sought most difficult. Nor did it help very much that the American team was most "permissive" in its surveillance of the progress of work in the collaborating countries. But this permissiveness is understandable because the scholarly community is such that one does not feel free to continuously inquire of one's collaborators as to how the research efforts are progressing in their countries. One is almost forced by the amenities and the proprieties of the "academic game" to extend unlimited trust to one's chosen collaborators, once the choice has been made, and once it seems as though there is agreement on the task and its various dimensions. Hence, one runs the risk, as we did, of encountering serious and basic misunderstandings at a time when it is virtually too late to correct them.

It must be reiterated that a personal visit was made by a member of the American team to try to ensure sound common understanding on the part of everyone involved as to what we were trying and what we wished to have done by our foreign collaborators. The possibility that some of the difficulty arose out of the problem of different languages does not seem very likely, since the language proficiencies of the American team were quite adequate to the task. It is true, however, that in two of the cases the personal contact of the American team member was with someone other than those who did the actual work on the comparative notation. That is, our personal contact and our enduring mail and cable contacts were with "senior" scholars who then delegated the task to junior collaborators. Somewhere in this three- or four-step flow of information there may have been significant slippage. So, too, somewhere in this stretched out chain of communication there must have been significant decline in the kind of personal scrutiny over the progress of affairs that now seems to us to be absolutely required for any such future ventures.

Perhaps these are overly elaborate efforts at explaining a failure in communication that may have been due to simpler causes, such as the failure on our part to be sufficiently explicit or the failure of our collaborators to understand or both. Perhaps, too, some of the responsibility for the tangential rather than direct contribution of the French and German teams is assignible to the extraordinary difficulties in phone communication that arose as a result of the earlier mentioned widespread social disorders that were characterizing France and Germany at the time we were in the middle of this effort. We were, for instance, unable to determine for several months whether our German collaborators had received our sets of instruction and materials.

3. To retrace several steps in the chronology of events -- after we had received reasonable assurance from several of our collaborators that they had pushed the work as far as it could be done at that time, we moved into the next phase of our collaboration, involving a ten-day conference at a European site. This conference had been planned at the inception of the international phase of collaboration. All of our collaborators were aware of the fact that we considered this conference crucial. Yet when conference time came, we had no word from our German colleagues and a last-minute message regarding their inability to attend from our French colleagues. It was not until several months later that we were able to reestablish contact with the German colleagues and discovered that though our communications had been received, they had not been read or had not been understood or both. They professed unawareness of the requirement of the conference itself. We find this difficult to comprehend, except for the fact that at the time our reminders were being sent, there was serious disruption of the university at which our German colleagues were located and it is eminently possible that our conference call may have been seen as something optional and, in any event, relatively trivial by comparison with the live events of the day. In any event, we were unable to take advantage of the person-to-person conference for the collection and clarification of data from our French and German colleagues. Our conference, then, consisted of one British colleague, two from Holland, and one from Italy, all of whom had presented documents in response to our requests. In addition, we were fortunate enough to have as continuing consultants two other colleagues, both Italian, but both specialists in sociology and education, and both with wide ranging knowledge about European education in general.

What We Learned from the International Comparisons

The principal gain from the collaboration with colleagues from five different countries consists of a fuller and more detailed appreciation of the difficulties involved in the evaluation of the effectiveness of educational systems. To the difficulties that we had encountered in applying various models and schemes to American materials must be added, we now know, a range of comparable difficulties that are sure to be encountered as one tries to apply any such model across lines of national systems.

The acquisition of such deeper and more detailed understanding represents an intellectual net gain, even though it is now clear to us that construction of a revised model, that might more adequately take international variations into account, is considerably more difficult than we had thought.

The comparative difficulties can be classified into several main types. These include problems arising from: 1) the different structure of school systems in the different cultures; 2) differences in educational goals; 3) differences in educational research; 4) differences in the relationships between the school system and the other publics and institutions of the society.

It will be remembered that we are speaking here primarily about distinctions between the United States, on the one hand, and England and Italy on the other, with only occasional reference to Holland, and almost no reference to France and Germany. This restriction in the comparative range was caused by different responsiveness of the collaborators from these countries to our request for comparative information. Yet even in the cases of the three countries on whom we do have come comparative information -- England, Italy, and Holland -- it became apparent, after a few days of exploration with our colleagues, that in varying degrees the orientation of the American school system, most particularly at the level of professed goals, differed sufficiently from those of the other countries (though least so in the case of England) to make it most unlikely that the basic purposes of the effort we had undertaken could be met.

It may be recalled from earlier chapters in this report that we had sought to avoid the problem of determining "whose" goals ought to be taken as the goals of education in the system by focusing on outcomes alone. The primary intent in shifting the focus was to secure an inventory of educational outcomes such that, whatever the goals might be and howsoever versions of these goals might be taken by various educational publics as adequate versions, an estimate could be made of the extent to which these goals were being met. It was also suggested that in the process of taking such an inventory it would be made clearer to concerned educational authorities how much in fact was reliably known or was not known about their respective educational systems. It was hoped that when the ratio of ignorance to knowledge, and the ratio of achievement to aspiration, were made clear by such an inventory, an indispensable first step would thereby be taken toward the development of an adequate instrument for the measurement of the effectiveness of national educational systems. For, we believed, and still do, that if educational systems, local or national, can focus selectively on those outcomes in which they can take pride, and can correlatively selectively ignore other outcomes or be indifferent to them and the lack of information about them, there is little chance for the arousal of educational conscience required for the development of sound evaluation instruments. Our proposed inventory-taking was designed to cut through the circle of self-reinforcing contentment with the status quo.

If we now ask whether our intentions have been served, the answer would have to be decisively negative. For a variety of reasons, we do not have enough sound information about the ratio of ignorance to knowledge or of achievement to aspiration, in our country, or in others, to "confront" educational authorities or ministries with the "facts of the matter" and thereby to arouse the kind of concern for more information, and for the reduction of the disparity between achievement and aspiration. Insofar as this kind of arousal seems to us still to be indispensable to educational reform, we must confess that our own efforts have not contributed much to that reform. Yet overall we are one step further ahead than we were before undertaking this project in the revamping of educational systems. For it is now clear -- or should be to anyone concerned -- that a much more elaborate and time-consuming effort is required if the indispensable first step of revealing the ratios of ignorance-to-information and achievement-to-aspiration is to be taken.

Some may argue that this exploration in apparent futility -- as this effort might from one point of view, be considered -- could have been made without the elaborate apparatus of dedicated research by an American team, complicated and compounded by the effort to include international comparative materials. One wishes that indeed had been the case. For there is little pleasure in engaging in such futile expeditions. The fact of the matter is that it was impossible to know beforehand -- before, that is, the kind of study that we undertook -- just what kinds of problems had to be solved and what kinds of difficulties would be encountered.

Nowhere, for instance, was it possible to determine before we undertook this research and the international comparative effort, how much information regarding educational outcomes was available in the various countries in which we were interested. Our expectation that this would require a serious effort at inquiry and discovery was based on our own experience with the American situation, in which we had immersed ourselves, only to discover that the problem of bringing order to the American data was gigantic indeed. We had little expectation then that the situation would be better in countries whose data collection technology on the one hand and whose explicit concern for educational evaluation on the other were not as far developed as in the United States.

Nor can it be said, with any serious attention to the facts, that simple inquiries to various national ministries would have yielded the information we desired, without the necessity of contracting with individual scholars to conduct such an inquiry. While it is true that there is more centralization of national educational efforts in all the European countries involved than in the United States, it is also true that neither these ministries nor the central federal authorities concerned with education in the United States are able to provide one with a systematic inventory of available information.

The reporting systems which have lately been developed in the United States (e.g. ERIC) in order to facilitate the kind of research we conducted serve as beginning guides to the range of possible information available. But they do not in any way replace or serve the purposes of an inventory of available information.

Moreover, our concern was with documented research findings regarding outcomes. At least, it was our intent to include only such information as met this criterion, however often in fact we departed from this criterion in the data we did include in our own inventory of U. S. information. None of the European ministries are in any position at all to provide authoritative information about the range of available research findings on educational outcomes.

Even the central educational reporting systems in the United States are unable to yield this information since these reports are restricted by and large to those educational researches which have been funded by one or another federal agency. While the bulk of large-scale educational research has undoubtedly been funded on a federal level -- at least in the last ten years -- there is a significant accumulation of educational research that has not been federally funded and that must be dug out of the various journals, chronicles, library files and books in which they are reported. The difficulties in using these data -- collected in so many diverse ways and with such diverse intentions and in such varying degrees of attentiveness to criteria of sound research -- have already been recited in the appended documents by Smith and Smith. So, too, the difficulties involved in using official reporting from local and state authorities have been recited at some length in the appended document by Kraus and Waldron. We do not wish to recite them here in further detail, but we do feel it important to call attention to the fact that the problems our own researchers found with using American data, official and other, seem to be even more oppressive in the European countries from which we recruited our collaborators.

On the basis of the discovery we have made of the inadequacy of data collection and analysis in all the countries concerned, there is no hesitancy on our part in asserting that there is vast ignorance in all the countries concerned regarding achievements of their educational systems.

Types of Ignorance

Non-Cognitive Outcomes

This ignorance is most profound in the areas of educational outcomes that fall outside the traditional concerns with certain narrowly defined cognitive abilities. Neither in the European countries nor in the United States is it possible to say with any assurance at all what it is, if anything, that school children have achieved by way of "cultural sensibilities" that would enable them to appreciate more sensitively and knowledgeably the "high culture" products of their respective countries.

The same observations must be made with regard to other non-cognitive areas of school accomplishment. No one can say, on the basis of available data, or at least as far as we know of these data, how much sound orientation to citizenship has been achieved by students as manifested, for instance, in their capacity to live more easily with persons who differ in ethnic, racial, and religious origins and proclivities. Nor is it possible to say whether the emotional health of school children is in any way appreciably different at the end of their school careers.

The area where there is most information is that of cognitive performance and here the evidence is best with regard to certain standardized levels of accomplishment on certain tasks involving verbal and mathematical abilities. Even here, however, the range of such abilities that are subject to testing, given existing technology, and that have been tested, given existing social interest in them, is narrow indeed. The mere mention of "creativity" as a dimension of the cognitive world of school children brings to light the scope and depth of existing ignorance.

The Role of the Schools in "Educational Outcomes"

If the school systems of all the countries involved are ignorant with regard to a range of educational outcomes, they are even more ignorant, it must be said, with regard to their respective roles in the production, maintenance, and elimination of this ignorance or of such knowledge as can be documented. The available information on what is actually known or capable of being done by children is far greater, such as it is, than the corresponding information on the extent to which the schools have been instrumental in these outcomes.

There is a bland assumption made in almost all testing of educational outcomes that the schools are primarily responsible. But this assumption, reasonable as it seems, creates certain difficulties. Thus, no one would argue that the amount of mathematics a child proves he knows on some standard test was probably learned at school (where else would most children learn geometry or algebra?). But the assurance of the important role of the school is diminished greatly when the subject in question is the capacity to use the language for communication, or the development of understanding of social processes, or the acquisition of insights into politics or history that the student may exhibit. The problem here is that the school tests focus primarily on the range of knowledge they reasonably expect a student to have acquired through his formal school experiences. They do not test the wider reaches of understanding or capacity that the student may have acquired elsewhere, through his own non-school, life experiences or, if in school, outside the formal curriculum.

Moreover, if the schools are responsible for the amounts of knowledge that children reveal on their tests, are the schools also responsible for the amount of ignorance that the children exhibit on these same tests? That is, can it not be said that by the same tokens that the schools pride themselves on the "accomplishments" and "successes" of their children, they ought also to take responsibility for the lack of accomplishment and the failure that every child exhibits to some degree or another on tests?

In none of the school systems on which we have available information is there much manifest systematic concern with this problem of responsibility of the school for the failures as well as the successes in children's manifest abilities or knowledge. Recent political events in urban areas in the United States have produced an ideology which allocated to the schools all the responsibility for failure and none for success. But this is simply the mirror image of the traditional stance taken by the schools in which they have tended to take responsibility for success but not for failure.

The political issue is not, however, central at the moment. It is the fact of ignorance regarding the efficacy of the school that is central. We have come to know this directly and indirectly. We know it directly from the sparsity of information derived from genuine evaluative research regarding the contribution of formal education to children's knowledge and understanding. And we know it indirectly from the sparsity of evaluations of outcomes of education that characterize the research literature. It stands to reason that if the schools do not know how much "cultural sensibility" their children manifest, they surely do not know to what extent such cultural sensibility was in any way cultivated or reduced by the schools themselves.

How Well Are the Schools Doing?

A third major area of ignorance which our explorations have revealed concerns the question of how well the schools are doing, where this question is to be answered primarily in terms of some previously established norm of expectation about output, relative to input. That is, if one asks whether the schools are doing as well as they ought to be doing, the answer is simply not available. The indifference to this kind of question is as widespread and profound as the ignorance regarding outcomes and the schools' roles in these outcomes.

We know of this lack of educational evaluation both directly from the research into the evaluational literature that we have conducted and indirectly from the absence of concern for such questions in that research. Moreover, we are reinforced in this judgment by our research into educational goals. For in this area we find a great diversity of philosophical positions regarding educational goals, but very little effort at operational statement of those goals, at stated costs, economic, social and psychological, that the schools are willing to entertain. Since it is not possible to say whether any enterprise is doing as well as it should be doing if there has not been a prior statement of the cost-profit arithmetic by which it proposes to guide itself, and since such cost-profit guides are markedly lacking from the literature on educational goals and outcomes, it is of course clear that this ignorance is characteristic of the field.

Why Do Some Schools and Some Students Do Better Than Others?

We have also come to be aware of the fact that there is a substantial dearth of research literature that might lead us closer to an understanding of why the students in some schools seem to perform better on certain standard tests than do others, on the average, and why in any given school, certain segments of the student body seem to do better than others on tested outcomes. Yet perhaps of all areas of research into educational outcomes, this area of differential achievement by various segments of the student body is most thoroughly under study, for it represents a central focus of interest of many of the newly emerging sociologists of education. The differentials between the averages on standard tests of the student bodies from various schools are only now beginning to be a matter of research interest, having been touched off (perhaps sealed off?) by the recent "Coleman" study of the impact of segregation and desegregation on school performance.

We do not intend here to continue to recite the areas of ignorance regarding educational outcomes and processes that are to be found characteristic of various national educational systems. Suffice it to say that with regard to the areas already cited, there is very little information to be had from any of the countries involved, except England and the United States, and even in these two countries, the research may be said to be just beginning.

It is an interesting commentary on the degree of similarity of the countries involved that wherever there is information that can be compared, it tends to be confined to the restricted areas of cognitive output -- verbal and mathematical abilities -- that were mentioned earlier. It is apparent that there is something resembling a common Western European culture system so far as educational orientations are concerned. Such variations as we find in the countries studied seems to have more to do with the differences between a remaining tradition of classical studies (as in Italy for instance) and the newer tradition of "modern studies" than it has to do with differences in the degree of a modern humanistic orientation. The distinction, in short, seems to be one between a balanced combination of classical and "pragmatic" studies, on the one hand, and an emphasis almost alone on pragmatic studies on the other. In general, however, all the systems seem above all to stress literacy with words and numbers as the core focus of their educational systems; everything else seems to be secondary. This can be seen in the reported outcomes from the Italian, British, and American systems (see appended reports).

It is only momentarily tempting to try to state this fact in some more precise numerical dimensions. Since we do have "counts" of the kinds of studied outcomes in these three countries, and since these concentrations of studied outcomes can therefore be stated in terms of percentage concentrations, it might be beguiling to give to this judgment a certain appearance of rigor. In fact, however, the data do not merit being treated as though they were rigorous. At best they make it possible to report impressions regarding the core concerns of the systems insofar as these may be inferred from what it is they test for and what they report. No other impression regarding these data commends itself so strongly as that regarding the concentration of the modern Western European educational systems on a type of education designed primarily to enable a person to secure a type of position in the division of labor.

From the exchanges on information at our conference, and from the papers submitted by our collaborators, it becomes possible to identify numerous differences among the various systems, around the core of common concern for verbal and mathematical literacy. These involve differences in the numbers of years of school presumed to be required to reach different levels of literacy; the kinds of other subjects expected of students; the number of years of school it is believed that all children should enjoy at the expense of the state; the percent of any age cohort who are deemed proper and suitable for advanced education; the relation between church and state; the role of parents in the governance, direct or indirect, of the educational systems; the degree of centralization of authority and commonality of curriculum through the regions and localities of the various countries; the amount of autonomy enjoyed by teachers; the kind of preparation teachers receive; -- all these and other differences are present and, of course, they are well known to any student of comparative educational systems. But these differences seem not to matter very much insofar as their influences upon the main thrust of the educational systems are concerned. For there the basic similarity in concentration upon mathematical and verbal literacy seems to override all other considerations.

We have come upon an apparent dilemma. For, on the one hand, we have urged in this report that the differences between the systems, as we discovered them, in such matters as structure, goals, research and school-public relationships made it most difficult for us to revise our model so that it might be suitable for the cross-cultural measurement of educational effectiveness that we sought to achieve. We added to this set of difficulties those arising from the lack of adequate information. But then we have most recently argued that each of the school systems concerned exhibits a central thrust that is basically alike in its concentration in mathematical and verbal literacy and this should make it possible for the model that we have developed to be applied across countries without any real difficulty -- assuming available information.

The real problem lies beneath the surface of this apparent contradiction. Our aim, it will be recalled, was to develope a model that when properly applied, would yield information about educational outcomes that might serve the purpose of revealing to the systems concerned how little they knew about their own operations, on the one hand, and how far away from their expressed goals they seem to be, given their own reported results. But as we have come to know the goals -- or the various sets of goals that might be imputed to each of these educational systems -- it has become clear that, except for England and the United States, there is little serious concern with the goals of education, except insofar as from time to time there may be modifications in official goal statements issued by the ministries, or modification in official declarations of intent to serve traditional goals more adequately in the future. Only in England and the United States does one find real ferment about educational goals, from within the ranks of official educational authorities themselves.

Yet, the bizarre facts of modern political history seem to have been caught up with all the educational systems and produced considerable educational ferment throughout the systems, but most particularly at the levels of higher education. This ferment has been brought about, of course, as is now well known, by the protest and dissent of students at universities throughout these societies. In all this protest and dissent, however, there is only the most meager indication of attention to the curriculum of the elementary and secondary schools. It may well be that more extensive research into these matters will reveal that in fact the dissent with university education has spread downward and is now being expressed with regard to the quality of education at the elementary and secondary levels as well. If so, then Holland, Italy, France, and Germany will come increasingly to resemble the United States and Britain, both of whom have been engaged in continuous and conflictful debates regarding elementary and secondary education since World War II, if not earlier.

It may also prove to be the case that the need to document the inadequacies of the educational systems of these countries, as a first step to induce educational reform, will be eliminated by the political pressures of students and others, who will assert and, if powerful enough, will enforce their notions regarding needed educational reforms. Political power, in short, may do what intellectual and scientific data might never be able to do, or would surely take much longer to do.

Yet it is doubtful that the kind of rational orientation to change, and the controlled experimentation with changed school practices that might have followed from intellectual proof of the need for change will result from very sweeping political demands. On the other hand, it may be that the atmosphere of change, generated by political activity of the kind we are now witnessing, may make certain bold experimentation far more possible than would have been imaginable under conditions of slow movement expectable from the pressures of scientific research findings.

Leaving aside the likelihood of change resulting from political dissent and pressure, it is evident that Britain and the United States differ from Holland, France, Germany, and Italy -- as far as we can tell at least -- in the extent to which there is a recognized disparity between expressed and claimed and desired goals of education on the one hand and the actual outcomes on the other. That is to say, both the United States and Britain, as educational systems, seem to be at least ideologically far more involved with a wide range of goals beyond those of cognitive skills than is true in the other countries. Hence, the utility of our fact-finding model, insofar as production of educational conscience and impulse to change is concerned, is likely to be far more limited in the cases of these other countries than in the British and American cases. For if there are no real expressed intentions to make education serve goals other than those traditionally tested, then the information that there is no information on these other outcomes is likely to have little or no influence. The absence of information can easily be attributed to the absence of concern. The appended report, for instance, from the French colleagues, regarding educational reform in France in the twentieth century, reveals unmistakably that concern for other than cognitive dimensions of education, and for larger numbers of school places for children, is minimal. But any reading of American or British documents on education reveals unmistakably that every public concerned with education seems at least as ideologically concerned about the quality and breadth of education as about the narrower domains of mathematics and verbal capacity that have traditionally dominated the center of interest of these systems in the past. It is because of this unequal political usefulness of an inventory of available knowledge about educational outcomes -- unequal because of differences in goals -- that we have come to have much less confidence in the value of our comparative model than we had before we undertook this enterprise.

The differences in the structures of participation in education decision-making among the involved countries also makes it likely that any such reporting on educational outcomes will be unequally valuable. It seems required, in the European countries, to make a dent upon the consciousness of national officials and bureaucracies before educational reform can be begun. But there is also much more coherence and integration between goals and purposes, on the one hand, and school activities, on the other, in those systems controlled by national ministries or bureaus. Hence, the likelihood of continuing self-confirmation of favorable estimates seems considerably greater than it does in a society such as the United States. Here, though the amount of federal intervention in school affairs increases greatly with every passing year, there is still fundamental local autonomy, such that it is possible to have numerous smaller centers of foment and dissent and experimentation which begin to create an atmosphere of needed change that may become contagious.

Moreover, it is possible for changes to be made, and substantial changes at that, in local systems, without involving the entire nation and without committing the stability of a national government to these experimentations. Thus, change may ensue and may spread, especially if facilitated by forward-looking national officials. On the other hand, it is evident that if European national ministries decide on educational reform, the entire system of the country is likely to experience the reform much more quickly than would be the case in a system such as the American with its emphasis on local autonomy. In any event, these differences in structure, and in the participating publics and the connections with the other basic institutions of the society make the proposed model for measurement of educational effectiveness much less useful for its intended ends in European than in American educational systems.

Finally, there is one important difference between the United States and the other systems that suggests that under some conditions the application of a measurement of educational effectiveness would be much more politically portentous and consequential in those countries than in the United States. It is the fact that, given the centralization and uniformity, relatively speaking, of education in European countries, and given the contrasting local autonomy and great heterogeneity and diversity in American systems, if European countries were to take seriously the question of educational goals and achievements, then any effort at measuring these achievements would likely be far more consequential for the entire system in quation that would be the case in the United States. Even if and when we secure a form of national assessment, as we are now in process of doing, the diversity of educational needs, as these are defined from one community to another, and the diversity of powerful interests who view the school systems' needs in quite different ways, are likely to be powerful deterrents upon any attempt to produce widespread change on the basis of the information provided by such a national assessment.

Yet one must put aside such hesitations and doubts about the value of information if one is to take education seriously as an enterprise and is to express any rational concern for the effectiveness with which educational resources are being utilized. This is true for centrally controlled as well as locally autonomous systems. It is true for traditional as well as modern systems. It applies to systems aimed at producing readiness for change as well as those aimed at producing predictable and steady citizens. It holds for systems concerned with the soul of man as well as those primarily focused on limited cognitive domains.

Each and all of such systems requires sounder knowledge about how well it is doing than it now has. And the question cannot be answered except as there is much more explicit concern than can now be found with the goals of education, and, in turn, with the development of some measures of outcome, and some ways of understanding the contribution of schools to these outcomes.

The analysis of the relationship between input, process and outcome is a matter which will take long years of patient research to refine to the point where it can become part and parcel of ongoing educational evaluation. Much work is afoot in that domain of research. It still remains true, however, that the impetus to develop such analysis and the impetus to innovate in education will develop -- all political considerations aside -- in proportion to the extent to which the school systems are informed about how well they are doing, given their intentions.

The first indispensable step is the measurement of outcomes -- in ways not now available. We now turn, therefore, to a brief consideration of alternative possible ways in which educational outputs might be conceptualized and measured, so as to provide a sounder inventory of outcomes than is now available to any system.

B. FINDINGS AND ANALYSIS

7. Alternative Models of Educational Systems and Their Evaluation

James H. McGregor

American education is the focus of intellectual power struggles -- struggles to define the goals of education, to specify those things which everyone must know; to determine just what the school is meant to do and then how to go about doing it. Courses seem to be added to curricula in some sensible sequence, but often explanations vary so widely, and expectations about the effects of a course are so great that no definitive statements can be made about anything.

American education clearly exists, people do go to school, some people learn something. But the avalanche of conflicting theories about what they learn and why demand at least codification if not integration. Changing demands on the school system, particularly the demand for real, not just formal, equality in education for blacks and whites place new demands on the curricula, theories of teaching, and ideas about what should come out of the schools.

An extensive effort is being made to organize and to evaluate all that seems to be going on here. Part of this program is the attempt to clarify and assess the goals and outputs of the total educational system. Analysts from a variety of disciplines have applied themselves to this task, and a variety of indices and supposed indicators have been constructed. The goal seems to be to approach in inclusiveness and reliability an indicator like the Gross National Product.

In the following paper, the gross national product is discussed -- first in terms of the sources of figures for the gross national product, i. e. the totals of all the monitored kinds of economic production. Then the underlying assumptions of the gross national product are discussed -- those features of a money economy which make this measurement possible. Clearly, if the gross national product is to be calculated, the economy must be organized in one of two ways: either there must be a money system, which is uniform in its methods of valuation throughout the system, or there must be calculable "terms of trade" in the system which enable an analyst to assign uniform value to all exchanged goods in the system.

In those systems of barter economy which do not have money, different methods of calculating gross national product are still available, although more complex. These barter economies have multiple products, and multiple exchanges, which, if the system is rational can be organized. Here the paper discusses very briefly some of the problems and methods of organizing a barter economy.

The purpose of both of these economic preludes is to present some way of looking at educational outputs, and hopefully to provide some framework for their evaluation.

We first examine the areas of educational output, delineating six general outputs of the educational system. Then, a system of units for each is examined. The purpose of this is to discover if some unit within each can be observed which is common to all. If this unit exists, then all educational outputs could be expressed in terms of it.

Since no such unit is found, the exchanges between these areas of output are examined. As if this were a barter economy, each output is examined, to see if there are uniform trades between one type of output and all others, or some others. If such a uniform, and rational trade were found, then we could calculate the terms of trade for the system of educational outputs, and on this basis develop a common unit for their expression. Since this, too, is unavailable, we examine the market mechanics of educational outputs.

We try to determine if there is any single market place within which all educational outputs, or at least some significant portion of them, are traded for something else. Therefore, we try to see where each kind of educational output is likely to be exchanged, and for what.

Here we discover several different markets: some trade-offs between outputs take place within the individual himself; others take place between him and his peers, or fellow workers, or his family and other intimates; some take place in formal occupational or legal contexts. Examining these exchanges we find that none of them yields either a limited marketplace for each kind of educational output, or one general marketplace for all educational outputs. It is at this point in the analysis that we conclude that without other information there can be no calculus of the gross national educational product.

The most serious block to the calculation of this product is found to be the three indistinguishable features of any output calculation. In every case where we seek to measure an educational output and are unable to do so, it is because we do not know whether:

1. the output is present or not;

2. the output is actually exchanged in this marketplace;

3. or, whether our methods of measurement are at fault.

We turn next to consideration of tests which could measure educational output in toto.

Several kinds of tests are discussed: the Scholastic Aptitude Tests and their usefulness in predicting the performance of high school students in colleges; the Federal Service Entrance Examination; and the general aptitude and intelligence tests used by the armed services. The strengths and weaknesses of these tests are discussed. It is concluded that the SAT is the most valuable model of this kind of predictive test, because it has a universe of past products with which to compare present applicants. There is a pool of successful undergraduates whose college academic records are known and whose test scores are known, with whom the applicant can be compared. His likely success in college is accurately predictable on the basis of the similarity between his test scores and the scores of other students in the college in the past. It is important to realize the limitations of this kind of testing -- one of which is the inability of the test to tell what the student actually learned in school and what he learned elsewhere.

Using these ideas of the sources of success in such tests as the SAT, the paper examines the proposed national assessment program in education. It discusses the limitations, and the biases of the test, and concludes that this test in no sense represents the kind of national assessment of educational product which would meet the demands of our economic analogues. It in no way compares with the gross national product in inclusiveness, or with any of the barter models in showing the tradeability of individual educational outputs.

The last section of the paper examines several perspectives on educational output and their associated measurements. It moves up a scale from the least general, but most measurable, to the most general and least measurable theories of educational product. The most important part of this analysis is its attention to the fact that measurability is a function of the structure of theory. That there are only certain methods of measurement which are consonant with certain kinds of theory; that the strengths and weaknesses of theory stem from the organization of the theory itself.

It is also extremely important to know in evaluating theories of educational output, or observational perspectives on educational output, that different kinds of assumptions are made not only about the ideal education in each, but that the nature of the educational product itself differs markedly in each.

It seems obvious that the product of the educational system is educated young people. This in fact is not the implication of many of the theories. Theories differ on whether the product of the educational system can be considered to be extrinsically associated with students, or intrinsic to them. This divergence is extremely important.

Extrinsic theories see the source of value in education in what can be done with it, and what can be done with those that possess it. They see the value of education as its value to society, either *qua* society or *qua* some institution (such as the economy) within the society.

Intrinsic theories of educational output see value determined in two ways -- they see it determined by *what the individual can do with it*, and they see it as valuable *in itself*. Thus educational output in this theoretical framework derives its exchange value from its intrinsic value, and it derives its intrinsic value from (eventually) definition.

The question of the value of education is an extremely important one, and one on which there is a great deal of theoretical fuzziness; most of which results from failure either to realize the necessity of making value judgments, or the mistaken belief that if education can be tied to some method of value assessment than its value will be "socially-objectively" discovered.

Tying the educational system into the money system is an effort to discover this socially objective value of education. By finding a dollar value for education, analysts involve education in a system of value calculation through exchange. Since everyone realizes that there are no objective values within the system of exchange, only values contingent on supply and demand -- it is assumed that this type of valuation will not be prejudicial to the complexity of educational output.

Similarly the attempt to tie educational output to those things which are useful is an attempt to objectify values of output. It must always be remembered that assertions of usefulness are assertions of value to the system, or to some part of it -- what is useful to the social system is valuable to it, hence it derives its usefulness from its value. If, however, the values of certain segments of the society are not examined, it is fundamentally unethical to posit usefulness as a guarantor of value either to the total society, or to the individual.

The last section of the paper then concerns itself with the relationships between theories of output, theories of measurement, and theories of value. There is particular care paid to the problem of separating measurement of value, from theory of value, and of carefully delineating the kind of output which is implied by a particular theory of output.

The "Total Product" of Education

The gross national product is the most comprehensive national assessment figure. In simple and easily understandable terms it tells how much the society has produced of goods and services during the period of one year. It does this by totalling how much money accrues to everyone as a result of the production of these goods and services; thus it includes:

1. wages and supplements (to individuals);
2. unincorporated net income (to unincorporated firms);
3. rents (to all owners of real estate and other immeubles);
4. interest (to investors);
5. dividends (to stock and bond holders);
6. undistributed profits (money made by business, held for reinvestment in own firm or for later distribution);
7. corporate taxes (the money made by government from corporations);
8. indirect business taxes (government income from this source);
9. depreciation (money paid out for the replacement of capital goods);
10. purchases from other firms (money paid by one firm to another).

The GNP then is a measure of production, and the value of that production within the society. Its existence is dependent on the existence of money and the fact that money can be exchanged for a variety of different things -- for land, for use of land, money, and labor, for goods produced, for raw materials. Money is useful in the same general way to everyone and has the same market value independent of who uses it.

In economies which do not have money there is no measurement of GNP possible. In a barter economy one can only tabulate the amount of stuff that is produced in a year -- so many potatoes, so many pigs, so many beads -- but unless all the trading terms within the society are known -- how many potatoes, how many pigs, how many beads -- there is little to say about the society. It can't be determined, for example, whether the society did better this year than last year. To know this without terms of trade, the society must produce more each year or less each year. There is no way to compare the case in which the society produced as many potatoes, fewer pigs, but more beads than last year. We can't really tell how good the market is, how effective it is either unless we know terms of trade, or have money. In an economy where terms of trade are known, if two potatoes can be traded for two pigs, and two pigs for two beads, we should expect that two potatoes could be traded for two beads. If there are no terms of trade, however, there is no guarantee that this will happen.

Calculations about the efficiency of production cannot be made either, without terms of trade. Imagine a wage economy: a man employs another to make beads for him -- in a year the employee produces four beads and is paid two potatoes; the employee trades one potato for one pig and the employer trades one bead for one pig. Is this a rational economy? From the employee's point of view he has traded four beads for two potatoes -- one potato, therefore, equals two beads. The employer pays one potato for two beads, but with his two beads he can get two pigs, whereas his employee for his two beads' worth of potatoes can get only one pig. Clearly the employee is selling himself short. Beads are overpriced; potatoes are overpriced; he is not maximizing his profits in this situation.

In short, then, if we do not know terms of trade, we cannot

1. measure gross output, only enumerate it;
2. compare (except rarely) outputs from year to year;
3. determine the efficiency of the barter system
4. calculate the rationality of any segment of the system.

An Intrinsic Unit of Output?

This brief economic introduction should help to elucidate the kinds of things needed before attempting to make some calculation of gross educational product.

We have seen that the gross national product is only calculable in a system with some actual medium of exchange, like money, or in a rational barter economy in which, although there is no unit of exchange, terms of trade permit assignment of unit values to commodities. We have also seen that commodity exchange values can only be assigned in a barter economy when the barter economy is "rational."

There are several things to find out about the educational system then, before we can say whether or not there exists a gross educational product.

First, what is the range of educational outputs? Secondly, is there already some composite of these outputs which represents an individual student's educational "product"? If no such meaningful product exists, is there possibly some system of exchanges between the different types of educational outputs which permit assignment of barter-unit values to each?

Finally, if none of these things exist -- if there is no existent unit, and no rational system of trade-offs -- is there some indicator, in the form of a commodity market where all of these things are offered for sale, which will enable us to place some individual and composite value on them?

The variable in our discussion will be: the different kinds of educational outputs, and the different marketplaces in which they are exchanged.

Grades tend to be thought of as some measurement of the output of the entire educational system. And to the extent that they can be said to be uniform, they do look like some measurement of what the school produces and some quality measurement about parts of that output. Some parts are rated A, some B, and so on. Do these show a correlation with any other kind of societal measurement? No, they do not. There is no association between high school grades and achievement in later life, in any of the ways this can now be measured. Grades do not correlate with an individual's life income, nor do they seem to discriminate between those more suicide-prone, or accident-prone, or those with high or low divorce rates. The variance in all these attributes is more complex -- they depend on social class, on family situation of the child, on religion, on marriage partner . . .

Is there a measurement which could be constructed which would give some indication, while the child is still in school, of the success of his school training, as this school training applied to his success in later life? It would have to contain, of course, all the areas of school interest:

1. physical wellbeing, and the child's ability to maintain this;
2. emotional wellbeing;
3. occupational preparedness;
4. basic skill preparedness -- English, mathematics, etc.;
5. moral preparedness;
6. drive, initiative, motivation;
7. culture-preferred knowledge.

Let us first discuss a barter model approach to this: by this we mean that the factors listed above are minimum specifications; they cannot be reduced to a single factor. There may be, however, sufficient trade information to work out a set of correspondences between them in order to do the latter. It is necessary to say how much physical health can be traded for how much emotional health can be traded for how much occupational preparedness, etc. Are there really observable trade-offs between these factors?

Take physical wellbeing, for example. The failure to maintain physical wellbeing produces death. Obviously there can be no trade-offs at this level, since death implies the inability to make use of other factors after its occurrence. Above this level, there are only trade-offs of factors implying physical health: trading sleep for increased study time; trading exercise and recreation for time on the job or for concert attendance. But immediately, these are more difficult and secondary calculations. To what extent is physical health itself diminished by each of these trade-offs and to what extent is something else maximized? They do not permit, without some consideration of the value of an attribute in relation to other attributes, to make value statements about it. There are, too, problems of different value assignments at different levels -- at the individual, small group, and at the national level.

We could hope to specify some necessary minimum level of physical health but even this would be only an approximation to a limit, and would not tell how to compare the value of other higher or lower amounts of health.

The same is true of mental health. To the extent that mental illness does not exclude the individual from participation in the society as a whole, and to the extent that mental health is a function of other than school-produced relationships or involves factors not considered by the school, no minimum values can be assigned to mental health. Assuming mental health could be monitored at the time of school graduation, and this is a different matter entirely, we would still be troubled by trade-off situations.

Occupational preparedness combines a variety of different factors. It can be considered the ability to meet the entry requirements for a variety of different jobs. Or, it can be considered to be the ability to perform any particular job: typing skills, clerical skills, training as an automotive mechanic, as a horticulturist . . .

Basic skill preparedness includes those things which are socially necessary: the ability to read a newspaper, to understand memoranda and other written directions, to understand mathematical skills and use them (at the level of filling out income tax forms and counting change in drug stores). It also includes the mathematical ability consistent with certain occupations, where these abilities are part of a general set of traits associated with general kinds of occupations.

Moral preparedness is an internalization, and willingness to practice most of the tenets of American law, and American group values. Obviously, very few trade-offs are permitted here at all.

Drive, initiative, and motivation contain two sorts of factors: general self-maximalization, in the sense of maximalizing one's own life-style and life changes -- preparing oneself for a better job and going out to find it, setting oneself up in a small business. It also includes approved innovativeness within occupations: the gas station attendant who, through efficiency and self-motivated hard work, attracts more customers to his station. Trade-offs are obviously implied here between other kinds of activities -- moral values, leisure, etc.

Culture-preferred knowledge -- this is perhaps the most poorly defined. At the least it includes some knowledge of literary and artistic classics.

Even this sketchy outlining of the general outputs of education elucidates many of the difficulties of national accounting and comparison. We have certainly seen that there is no unit intrinsic within each output which permits us to compare them. It should also be clear that we cannot, as we perhaps can in our barter model, develop "terms of trade" for them by examining their exchange values with each other. We simply cannot imagine them being exchanged. There are systematic reasons for this difficulty of exchange, also, which we can note.

1. The factors have different loci -- they pertain to different parts of people or of the same person: does an individual have any way to decide to sacrifice a certain portion of his mental health in the interest of increased knowledge of the classics?

2. The factors have different minima.

3. The range of each factor is poorly specified -- the difference between drive and initiative and moral preparedness is not clear. What is really initiative and what is normatively prescribed initiative?

4. The factors are often inseparable -- physical and mental health overlap, and other factors overlap with each. Thus, exchanges can not always be seen because the categories are not mutually exclusive.

5. The lack of mutual exclusiveness is further compounded by difficulties of surveillance and observation of incidence.

A barter model search for output unit

The difficulties of surveillance are actually the critical point of the entire analysis in a way. Just as there are a variety of factors to include, so there are a variety of observers, assessors, and ultimately purchasers of these capacities.

In what market, then, are each of the factors exchanged?

1. Physical health -- is bartered only within the individual. He may barter it in exchange for factors available on some other market, say, the occupational market, but the actual exchange is an internal one. There are considerable differences in assessment of the individual worth of physical health vs. some other gain. Thus there is little likelihood of any overall market value placed on physical health.

Imagine a situation for example in which two men enter into competition for the same job. Man A trades two hours of sleep a night for extra study time to prepare for the competition; man B loses no sleep because of his relatively free evening schedule, but because of the increased anxiety develops an ulcer condition while studying for the examination. Now supposedly these factors could be compared: a certain increase in occupational preparedness is worth loss of "x" amount of sleep is also worth a predisposition to an ulcer, hence "x" amount of sleep loss equals predisposition to an ulcer. This would be marvelous if we knew that both were rational producers; that in fact neither had exceeded the minimum cost in preparing for the examination. We would also need to assume that at the beginning and end of preparation both were at correspondent levels of preparation and that both had received the same grade on the exam. If in preparation for this same exam a man C had lost no sleep and developed no ulcer predisposition and achieved as well as men A and B on the test, we would be at a complete loss to assign exchange values to any of the losses of physical health. But, at the same time, can we dispose of the problem by postulating health as a minimum which must be attained and maintained by all, because it is quite reasonable to assume that all individuals felt equally satisfied with the conditions imposed by the test. Health is tradeable, but certainly in no uniform way. Nor is trade always observable.

2. Emotional health could be treated in exactly the same way.

3. The marketplace for occupational preparedness is the actual job market. An assessment of the qualifications of the prospective worker is made by the hirer and on this basis a job is assigned to him, or he to a training program, or he gains entrance to some qualifying examination. Here we have the questions at a somewhat different level, but they

are the basic questions: a) what is exchanged and b) what is paid out in this exchange?

a. the applicant can be bartering his occupational preparedness, but he is also bartering --

b. his expectation of getting a better job

c. the expectation of higher present gains in salary for a job with more possibilities for advancement

d. some portion of his time

In exchange, he receives:

a. a contract for a certain amount of time; or

b. a training program; or

c. a program with limited advancement but high salary now; or

d. a program with high attrition rate, but high advancement.

We are simply uncertain about what is exchanged. There are several possibilities on both sides for calculating the value of the exchange. We are uncertain about the marketplace. How much is exchanged within the individual -- his exchanges between present and future -- and how much of it takes place between prospective employer and employee.

4. Basic skill preparedness -- these are the factors with which the individual deals with the empirical elements of his universe, mathematically, linguistically, and socially (including such things as relevant "social sterotypic knowledge, where to go for what, the details of how to get it once there, and what to do with it afterwards).

Let us call it "poise and preparedness." How much of this does an individual need in order not to be inconvenienced in his everyday social relationships? How much linguistic ability does he need, in order not to be considered a total fool, or to be prejudged invidiously in social exchange? To what extent does linguistic ability or inability condition the kind of job he will be considered acceptable for? If he can't speak English well, will he be made office manager, even when he is the most skilled and knowledgeable clerk? There are an infinite variety of exchange points, or markets, here:

a. within the individual in terms of developing poise and preparedness;

b. in interpersonal relationships, friendship, etc;

c. in formalized exchanges on and off the job (will others think well of him, how does he evaluate the importance of others' opinions of him?);

d. in economic exchanges (is he an intelligent consumer, does he lose money because of inability to count change or shop wisely?);

In all of these different markets, basic poise is evaluated and rewarded, the amount present in the individual is called into play, and some return is made to the individual.

5. Moral preparedness -- there are several formal and informal observances of moral preparedness:

a. the legal structure: police, courts, judges;

b. the occupational structure (what is the legal record of the applicant? Informally, the employer questions basic trustworthiness, honesty);

c. informal extraoccupational moral assessments -- credit ratings, for example. Ability to trust between friends and associates of all sorts.

6. Drive, initiative, motivation -- this too has several areas of exchange:

a. formal occupational assessment of the capacity of an employee to serve beyond the literal demands of his job;

b. the exchange valuations of "fellow worker" vs. "good worker" made in many informal occupational settings. These are exchanges which both the individual and his fellow-workers make. The possibility of exchange between value of initiative to the individual, assumed value to the employer, actual value to the employer, function or dysfuction in peer relations;

c. in intra-group relationships, the role of group leader, or opinion leader, who picks activities formally or informally, who is a deacon of the church vs. being a loveable fellow.

7. Culture-preferred knowledge -- this may "come out" as the ability to participate in intellectual discussions. It may be simply the ability to find enjoyment in reading in one's leisure time. To the extent that it is purely individual, it represents an exchange made within the individual between other factors like conversation, or television, or going for a walk. There is little chance for articulated social exchange.

Summary

We can see, then, that there is no common marketplace on the basis of whose rules we can assign exchange values to all of these factors. Yet we can also see that they cannot be uniformally assigned to such things as the individual, the economic marketplace, friendship roles, etc. In essence then we have posed some features which are indistinguishable from the point of view of the total system, and from the point of view of various actors within it.

To summarize, the reasons for these indistinguishable factors:

1. we cannot show that each area of output is measured in any kind of exchange in only one marketplace;

2. nor can we show that any single output is ever evaluated in any single marketplace;

3. we cannot show significant lumpings of outputs in such a way as to say these three things are always associated, and we cannot judge them independently, but we can judge them independently for the following items;

4. we cannot show that all outputs are evaluated in any single marketplace. We are at pains to show that any single exchange does not implicate at least most other factors. As a result of these considerations, we can say that the following three explanations can be applied in every case to educational outputs:

a. our measurements are not correct; or

b. the output in fact is not present in the desired amount within the individual; or

c. the output is not in fact considered in the particular exchange studied.

Conclusion

We have shown, then, that:

1. there is no unit common to all educational output areas which enables us to quantify them;
2. there is no uniform exchange value between educational outputs, which would serve as "terms of trade" and enable us to calculate the total product;
3. there is no single marketplace with rules which enable us to assign uniform exchange values to educational outputs.

Tests of Output

Since we cannot by observation or exchange analysis set some value on educational output, perhaps we can construct some test as a measure of educational output, which will enable us to assess total educational output.

Ideally this test would be constructed so that many things could be learned from it.

1. the amount of each kind of educational output which is present in each graduating student;

2. through some rule of additivity we should be able to calculate the total amount of single output which is present in any group of graduates, and the total of all outputs present in a group;

3. we should be certain that a measured output is actually one produced by the school system;

4. we should know the correlation between this output and other factors associated with adult life. We should know how well these factors correlate with occupational success, mental health, marital stability, etc. In other words, whenever the school makes a contention about the applicability of its output, the test should tell us whether the output is in fact produced, and whether it is applicable in the ways it is said to be;

5. we could rephrase point four, by saying that the test should be predictive.

At the moment there are not tests of this sort. There are simply no measurements of educational output which answers all these needs. There are several kinds of tests which are given to wide samples of people. Some of them are highly predictive of the things which the test-makers want to know. We will examine some of these and on the basis of their successes try to construct some idea of a suitable assessment tool for education.

Finally, we will examine the national assessment program as envisioned by Educational Testing Services, and discuss its adequacy as such an instrument.

Existing Models

Most universities have elaborate systems of computing how well a student will perform within the college on the basis of high school grades and on the basis of national aptitude and achievement tests. The Scholastic Aptitude Test has been applied with increasing scope and effectiveness and has had several major effects. First, it allows schools with some surety to expand their sources of admission beyond their traditional boundaries without high risk of lowering academic requirements. It is possible for admissions offices to make fairly safe projections (with a specifiable error range) about students from high schools with which they have had only limited experience, on the basis of combined knowledge of high school grades, SAT scores, and recommendations, perhaps combined with an interview of the candidate himself. The total so computed, while not measuring the output of the school in any sense, does create a profile of the student. The consonance of this profile with profiles of previous successful admittees permits the school to make its choices.

A similar nationally available test is that for entrance into the Federal government: the Federal Service Entrance Examination. Although the figures are not available, since the test continues to be used and refined as a selection instrument for federal employees, presumably there is some correlation between performance on this test and performance in the federal bureaucracy.

The Army uses a variety of tests first as discriminators of those fit for induction and then to assign inductees to various programs within the military.

It should be clear from consideration of these three tests that they are unlike the kind of assessment program we have been talking about in a variety of ways. First, they do not attempt to measure individual outputs of the school system in strictly limited ways. They are not batteries of tests which measure "drive, initiative, motivation," with one instrument, and "cultural knowledge" with another. They do not presume to assign numbers of these characteristics. Instead they are intermediaries between an unknown something -- the education and "knowledge" of the individual and a relatively known something else -- what their present and past memberships have been like.

These tests do not discriminate, as we must, between those things which a man learned in school, those things which he learned at home, those which he picked up in conversation before the exam. This is obviously unsatisfactory as a measurement of school output, because there is not even an imputed association between what the testee does on the test and what he learned in school.

Secondly, the test succeeds because there is a universe of desired traits. Yale University can admit high school students, not because they have selected out the "best" from among them, but because they have selected students who look the way other Yale University undergraduates have looked in the past. It is noteworthy, too, that the test is not used to predict specific achievements -- this student will receive the following grades in English, these in math -- but rather they predict overall grade-point average at Yale.

One should be rather more suspicious of the claims of the FSEE and the Army tests to predict performance, because of the evaluation of those who in the past succeeded on the test. Since the FSEE prepares for federal hiring, and since civil service limits the ability to fire people once hired, there is little chance that any but the grossest errors in prediction could be discovered. The test may, in fact, discriminate very little between those who are to be hired and those who are not. It is quite possible that among those who are hired there is as random a distribution of performance capacities as there is among the unhired. Here the test is even less predictive than the SAT as used. It tells the tester how well this testee compares with the universe of all who in the past have taken the test. It does not tell him how well those people have performed their jobs. Nor does it discriminate, as does the SAT, between those who will perform at a high level within the acceptable range and those who will perform at a medium or low level within this range.

Similar criticism can be made of the Army tests. First, because as is the case with all the tests there is little effort to find out about those who fail the test. Nor is there any evident way to find out about the range of achievement of those who pass the test.

The National Assessment Program

Data on the National Assessment Program, although plentiful on the intended content and scope of the test itself, are very limited in their attention to just how the test is to be used.

A brief introduction to the booklet "Phase I: National Assessment," states a threefold purpose of the test:

1. "It would give the nation as a whole data on the strengths and weaknesses of the American educational system. Thus it might constitute a much more accurate guide than we currently possess to the allocation of public and private funds -- where they are needed, what they achieve -- and to many other decisions affecting education.

2. "Second, assessment results, especially if coupled with auxiliary information on characteristics of various regions, communities, schools, etc. would provide data necessary for research on educational problems and processes which cannot be undertaken now.

3. "When sampling and testing procedures are adequately developed, international comparisons might be possible." 1

Sub-goals include "mak[ing] all groups more interested in the educational system . . . ." It is stressed that the test results should be intelligible to the layman.

The test is to be given to four age groups: 9, 13, 17, and 30 year olds. Three of these groups will be within the schools and a fourth will not.

The first purpose of the test is to give data on the strengths and weaknesses of the American educational system. The limits on this data are heavy:

1. the data can only be compared between age groups, at least initially;
2. there is no certain connection between many of the items tested and the curricula of the schools in which the children are students;
3. the tests may be "context-bound";
4. the tests are not predictive;
5. the tests are not additive at an individual level.

1. Comparison between age groups: the theory of the test is based on "tem sampling"; it tests randomly certain items from some larger universe of possible items. The example of a dictionary test is given for consideration; here the tester is justified in taking a random sample of words from the dictionary and, on the basis of percentage score in this random sampling,calculating what corresponding percentage of the entire dictionary of words the student knows. This is a relatively simple example, however; one in which the full universe of items knowable can be specified. In these particular tests, however, there is no "given" agreed universe. The test is a random sampling of items like those items on the test itself. It is assumed that the universe is very large.

If we confine ourselves to one test of this universe, there is no difficulty. Thus, if we confine ourselves to comparisons between a single age group, we are in fact constructing a hypothetical universe of things like the things on the test. If, however, we try to say that there are various levels of knowledge within this universe we may be misled. Clearly in the example of the dictionary, we would not. We could construct a test using item samples of the easier words in the dictionary, and another test using item samples of more difficult words in the dictionary, for use with an older group. However, we are not here dealing with this same kind of known universe, and it is quite possible to suggest that the item universe implied by the test at age nine is quite a different one from that implied in the test at age 13, and that at age 17.

Clearly, to get an accurate measurement of the range of abilities at age nine, and a similarly accurate range at age 13, the test must ask different questions of the two groups, or propose different tasks. The question is simply whether the tasks set to the two age groups correspond in such a way that the tests measure the same field of knowledge -- but one measures the amount that should be measured in age nine children and the other measures what should be measured in age 13 children. Comparisons of age nine vs. age 30 present greater uncertainty about the continuity of the implied universe.

2. Even within the age groups we do not know just what the test has to do with the product of any school or curriculum. Indeed, every effort is made in the tests to tie the content of the test with the desired programs of the schools. But anyone familiar with the diversity of goals in American schools will realize the often nebulous relations between goals in different schools. The test tries to approximate most of these but this does not account for several sources of variance within schools: non-overlapping of goals, and the problem of differently spaced curricula -- some schools do the same things in different orders.

There are serious interpretive problems in comparing scores between schools; we can predict the following general results:

Scores in the north and east will be generally higher than those particularly in the south, but also those of the mid-west. Scores in schools with higher concentrations of middle-class children will be higher than schools with more working-class children. White schools will score higher than Negro schools.

It will be most fruitful to compare scores on tests between schools which are relatively the same on all external criteria. It will only be in such comparisons, with external factors of *race*, *class*, *region*, *education of parents*, etc., held equal that schools will be able to say whether they *as schools* have done or not done their jobs properly. In any other kind of score comparison there will continue to be debate about the effect of class, race, etc. Should the test show, for example, that white suburban middle-class school children do better on all scores than lower class inner city Negro children this will not predict what is required to alleviate these problems. As we have said the test may discriminate for better teaching and better education within comparable social environments but it cannot discriminate between learning in school and learning outside of school.

3. Lord, who is the author of the text on methodology cited in the booklet on national assessment, states that there is no way to tell whether or not results on the test are context-bound.[2] This is to say, whether or not the results of the test are generalizable or whether they refer only to ability to perform within this specific test environment.

4. The predictive nature of the test is of great importance. In the example of the college admissions test, prediction of college performance depended on consonance between the individual profile of the testee and the profiles of successful college students. There is no profile of comparison which can be used for the present national assessment test. Indeed the test is simply an inventory of some areas of knowledge without any information at all about what this knowledge means; what it enables one to do; what its relation to likely performance in the future is; or how necessary it is to have this knowledge.

Perhaps the scores of 9, 13, and 17 year olds could be compared with that of the 30 year age group. Other tests could be devised of the occupational success, the personal happiness, the civic status, the income--of the 30 year old population; correlated with their assessment scores; and compared with the 17 year old population. This, however, involves an almost ludicrous assumption. If changes in school policy are to be made on the basis of comparisons between 17 year old population and 30 year old population, this assumes that the criteria measured in the 30 year old population are the ones which cause their happiness, occupational success, income, etc. If not, then structuring schools after the 30 year old population is rather like painting apples orange and assuming this will increase their citrus content. We are not exempted by the existence of 30 year age group from inquiring what is responsible for their success and happiness, what did they learn in school that they used throughout their lives, which enabled them to get better jobs, which enabled them to use their leisure time constructively, which prompted them to learn about art and music and to attend symphony concerts themselves. Obviously the only thing which comparison between age 17 children and age 30 adults will tell us is how similar they are on a series of measure indices; whether these indices have any causal correlation with life-style and life-chances, this test cannot tell us.

Differences between the scores of 30 year olds and 17 year olds can be explained in the following ways:

a. the schools have changed -- either in emphasis or effectiveness;

b. the school populations have changed. (Negroes are probably somewhat better educated in the 17 year old population than they are in the 30 year old population);

c. scores on science in the 17 year old population will probably be higher because of increasing attention to it in the late 1950s. Mathematics scores in the 17 year old population may be better because of changing math curricula. The number of schools offering art and music courses has also increased since the early 1950s;

d. scores of 30 year olds in some areas will be higher since the 17 year old population contains no college graduates;

e. differences in scores may simply be a function of the fact that the test is context-bound, and whereas 17 year olds are used to taking tests, it has probably been some time since 30 year olds took any tests at all.

f. people forget things between age 17 and age 30.

What will these things tell us about the school at present? If, for example, we find that almost no one in the 30 year old population has any knowledge of art or music, or any real critical capacity in these areas, should we stop teaching art and music in the schools? If we find that only some professionals retain high performance levels in mathematics, should we stop introducing higher level mathematics courses into the schools?

The Effectiveness of National Assessment

In short, then, all we have seen that the national assessment program will enable us to do is to compare schools in similar social settings. It may serve as a channel for research funds, and this will be of some value, in isolating those schools which are best to study. It will not, however, prevent the continual debate between those who take the view that ghetto schools must be improved before ghetto living conditions can be improved, and those who argue that unless the ghetto is changed there can be no improvement in its schools.

The test, then, simply will not answer, nor should we expect it to answer any of the "what" questions. If can within a limited range say "how well." It is no magic answer to any of the things which the school needs to know about the relevance of its curriculum to preparing the children for post-school life. It may tell some of the problems which exist between schools; why high school students can't read adequately, for example, may be a function of how poorly they read in junior high school. Some of the in-system inconsistencies will be eliminated this way, doubtless. But basic questions about the relationship between the school and the external society simply are not answered. Research still needs to be done on how certain groups of people get jobs, and others do not; how some perform well in school and others do not.

The Variety of Educational Outputs

In society there are a variety of people who are concerned with educational output, and there are a variety of "uses" of education. It is useful for getting a job, for watching a play, for raising children, for making oneself happy -- educational theorists, educators, employers, and educated people all contend that these are ways in which education is important.

When we evaluate the achievement of the educational system, then, we want to keep all these different uses of education in mind, and we want to keep all the different kinds of people who evaluate education in mind. We should construct a theory of educational output which will enable each of these people to say how well the American educational system does what he wants it to. (Clearly there are some people who will be disappointed -- the American educational system is not allowed, for example, to teach people the theology of any particular religion. People who want an answer to this kind of question can receive none.) People who want to know how well education prepares one for college, how well it prepares one for a job, whether education is making people more happy, more able to do what they want, more intelligent in their use of leisure time -- people with all of these interests in the educational system have almost no answers. It should be the task of an educational analysis to provide answers to as many questions of this sort as it can.

It should also be the task of analysis to suggest ways in which each of these outputs can be improved. Or if any area of output is seen as insufficient, there should be ways to suggest changes with some certainty of success. We want to know too whether the educational system is imbalanced in its attention -- is it devoting too much time to occupational preparedness, and too little to preparedness for intelligent citizenship?

In essence, then, there are a variety of observers and evaluators of the success of education, and there are a variety of different theories structuring the educational system. In order to evaluate the educational system most effectively we should know, not only what the output of the educational system is, as seen from the perspective of one of the groups of people to whom this output is of interest, but to all different interests in education.

Particularly when there are competing or overlapping foci of attention on educational output it is important to have a standard unit of measurement. If we wish to know what creativity has to do with leisure time, and what creativity has to do with getting a job, we are forced to use two different systems of evaluation, those which center on the structure of individual values and those which talk about economic uses and rewards of creativity. Indeed, anytime we wish to evaluate the total effect of some area of school output, to the extent that this is observed by different "people" or in different "markets," we need to have translation units from one sort of evaluation to another.

This returns us to earlier statements about the importance of translatability in educational output. It is useful, but not ultimately useful, to treat educational outputs from a variety of inconsistent perspectives. From the perspective of occupational evaluation, from the standpoint of individual happiness. But since educational outputs are eventually vested in people, and since people do a variety of things with those outputs, it would be well to have some system of translations from one to another. And in considering all the areas important to one educational output, such as creativity, or writing ability, there should be some way to examine situations where some skill or attribute has a variety of uses and to assign some quantitative value to the amount present within each student so that the full range of his potential can be examined and known.

Orientations of Observation

There are a variety of perspectives, then, which must be considered, some are more inclusive than others, some give more exact measurements within their ranges of attention than others. They are:

1. cost/efficiency models
2. resource developmental models
3. social models -- integration-disintegration dimensions

   social change model -- mobility-equality dimensions
4. psychological models -- personal focus; additive personal focus
5. humanistic models -- man in the universe (religious) focus; man in western society (cultural) focus; man in the modern state (citizenship) focus

Since cost-efficiency analysis pertains to all other forms of analysis, we will consider it as part of each system of analysis discussed.

Resource developmental models treat the child as a "carrier" of talents. These talents can be of many kinds, but they are all very much the sort of "talent" that a music teacher might like to have in a star pupil--a natural affinity for the instrument. Or it may mean a certain kind of analytical mind, which makes the child "ideal" for a career in math or science.

There are three kinds of resource developmental models: a) one in which all abilities are seen as talents and the task of the school is to maximalize any for which a student shows ability; b) another in which only high levels of ability are developed; and c) a third in which certain relatively rare capacities are sought and developed.

The first model simply seeks to maximize the ability of each student. A wide variety of courses is offered by the school and each child eventually channels himself into one area of courses -- the school makes no special effort to direct this choice and does not concentrate on offering any specialized range of courses. This model is so loose and so unpredictive, however, that it is little more than a description of the breadth of school offerings. It is indistinguishable from several of the humanistic models which we will treat, and similar to any socialization model of education. The major difference we will see between humanist models and this particular resource developmental model are the normative provisions of the humanist model. The resource maximalization model is undirected; it posits no area of concentration better than any other; in fact it has no real criteria for making choices about the range of curriculum.

As soon as we impose criteria for making choices about curriculum, within a resource developmental context, then we have to use one of the two remaining resource developmental models. The school delineates some areas of societal relevance, e.g. certain kinds of occupational preparation, and centers its course structure around these.

There are thus two major kinds of resource developmental models, one in which every child is given access to every kind of training; most excellent students in each area are identified and given more difficult tasks and increasingly higher levels of training in whatever area.

The second is a more limited sort of resource developmental model. It focuses on the search for certain relatively rare capacities and sees the exploration of these capacities by all children as a necessary means of finding the most able. Much of the post-1958 math and science curricula of American high schools is conditioned by this latter model. All children are exposed to science and math education, partially because this sort of education is of value to all children, but primarily so that these societally useful talents can be made visible; so that talented children can discover an interest in these subjects, be encouraged to pursue this kind of training in college, and these kinds of careers in life. The percentage of able children so identified is likely to be quite a small portion of the student population.

This is as clearly a theory of goal organizations as it is one of outputs. When combined with measurement methods this theory is one most likely to be applied by economic analysts of education. The method of measurement depends on the ability to isolate and observe these special products of the system -- the advanced students in a particular subject.

It is quite easy to make a cost-efficiency analysis. One simply looks at the cost of all courses which can only be considered to be of the type useful only to "talented" students, e.g. advanced mathematics or physics courses in high school, and add up their cost. Next one adds up the number of physicists and mathematicians, specifying the level of skill desired, eventually produced by any high school graduation cohort and by division calculates the cost of producing one such scientist. This cost compared with his life-time output (this cost further include college and post-college costs) will determine the productivity of finding him and training him. His productivity can be computed in a variety of ways -- the total productivity of all pure scientific research averaged, or one can simply say his "market price," expressed as the life-time salary of a scientist.

This is a very simple sort of analysis, it includes a number of assumptions about the nature of a society and the nature of goals and value in that society. For one, we have directed the entire school system towards the production of an elite possessing certain characteristics. We have also evaluated the "worth" of that elite entirely in terms of its monetary value.

With these implications in mind, let us turn to a similar model. This is an industry model of education. The analysis is one undertaken by Machlup -- he attempts to measure the output of education -- which he calls "knowledge production."[3] He also identifies "knowledge producing industries." The purpose of education is to staff such "industries" and to provide personnel in isolated occupations who produce knowledge. (Knowledge is so loosely defined, however, as to be almost a question of individual choice.) He then measures the productivity of these knowledge-producing industries. He calculates the actual percentage of the cost of education which is directed towards the production of knowledge, and then adds to it the cost of other forms of education (education in the home, calculated as the income lost by women who do not work but stay home and take care of pre-school children). This gives the cost of education in economic terms. He then calculates the amount of money paid for the product of the knowledge-producing industries, and observes the productivity of the industry, which is simply the ratio of output/cost.

What about the people not in knowledge-producing industries, or in knowledge-producing occupations? What does this evaluation have to say about them? Virtually nothing. In fact just as the resource development model which we discussed, treated non-scientists, so this model treats other students as unavoidable expenses in finding the people who really need training. They spend their lives doing whatever it is that they do, and are the subject of analyses of different kinds of economic observers.

There is, of course, a very sensible reason for this from an economic point of view. If the aim of the analysis is to discover to what extent the educational system contributes to the gross national product, it is necessary to differentiate that population. In a system with universal education, where all workers are educated (supposedly, and increasingly so), it makes no sense in terms of economic analysis to calculate the productivity of the entire labor force as if all their productivity were directly the result of education. We deal with one industry, the knowledge-producing industry, and call this the "output" of the educational system. Then we can make some kind of calculation and avoid the meaningless proposition that the entire national product is the output of the educational system.

This form of analysis grew out of the necessity of calculating the effectiveness of education in raising the GNP of underdeveloped countries. It was developed primarily by Harbison.[4] In this setting it is extremely important to know the amount of GNP produced by education itself.

From the standpoint of the goals of education as these are most often imagined, these analyses are unsatisfactory. Their inadequacy, of course, lies in their inability to treat the full dimension of educational outputs. Attempts to tie the whole range of school goals to this kind of analysis are extremely unsatisfactory, but they have been made.

The simplest of these analyses attempts to state all the factors which can be of relevance in occupational performance, and implicate them somehow in the economic rewards for performance. They delineate several characteristics of a good workman or a good executive: initiative, ability to get on with others, leadership ability, normative compliance, language skills, mathematic skills, willingness to learn, creativity. They attempt to assess the importance of these factors in the performance of a job, and then on these bases attempt to assign a dollar meaning to each of these qualities.

Resource Developmental Models: Summary

All of these models, then, with varying degrees of inclusiveness treat the output of the school in economic terms. We have moved up a scale of inclusiveness, from the minimal resource developmental model, which treats only one kind of output (mathematicians and scientists), to a more advanced resource developmental model which treats scarce personnel in all areas. We then discussed an industry-output model, which treats all education as if its sole production were "knowledge-producing" people and alluded to the kind of model which attempts to assess all the output of education in terms of stratified skilled people.

Excluding the resource maximalization model, which is not really a predictive model, there are four general characteristics of resource developmental models:

1. the school output is seen as only a part of the graduating population;
2. the school output is assessed only in terms of one societal measurement;
3. the school output is intermediate between what goes into the school, and what goes into the society;
4. the school output is "useful" to some real segment of the external society.

These four statements describe the strengths and the weaknesses of the resource developmental model. They are principally dependent on the fact that the resource developmental models all see their outputs in economic terms. Thus, they have a unit of expression and comparison -- money, which unites them to other things in the society, and which gives them a value which can be expressed and manipulated in the same ways that all similar values are expressed.

By expressing these outputs in monetary terms the analyst can:

1. compare educational expenditures with other kinds of expenditures in the society;
2. if outputs can be differentiated, and separately evaluated, assess the productivity of different kinds of educational output, and compare them.

However, the weakness of resource developmental models comes from exactly those things which give them their peculiar strength. As we have seen they have an objectivity produced by their relation with other kinds of societal measurements through money. This interrelatedness, however, prevents their being treated as intrinsic. By this calculus, for example, we cannot say what the value of education is to the student himself. How productive is it of happiness, of wellbeing, of comfort, of excitement? In order to tie educational outputs to some real societal measurement, they have been made extrinsic, i. e., they exist to serve some other purpose. Although they may have value in themselves, there is no way to measure this value.

This is by no means the only kind of theory of educational output there can be; in many societies educational output is not extrinsic at all, but totally intrinsic. The kind of socialization into adult roles which we find in most societies is exactly like this. People are taught to be adults -- after they have learned to be adults they simply practice being adults -- adulthood serves no other purpose.

We have seen that in resource developmental models educational output is good only insofar as it contributes to individual or national wealth. Wealth and productivity are seen as the supreme good, and education serves this end. It is important here to note that what starts out to be a measurement device -- efficiency and productivity -- in the absence of considerations of efficiency for what and productivity of what becomes a normative evaluation of the educational system.

An example should help clarify what this means. We cannot evaluate in any of these forms the importance of helping children to read the Odyssey unless this reading contributes to the way in which the child will eventually earn his living. This is not to say that he must become an English teacher in order for this skill to be valuable; it must, however, somehow be involved in the part of his life which is evaluated and rewarded by the money system. He must be able to take what he learned from reading the Odyssey, or what he became as a result of that reading and sell it on the money market. The same is true of such things as his ability to love birds, or his knowledge of biochemistry. Even more interesting and perhaps somewhat more surprising, such courses as home economics for girls, whose practical value should be obvious, must be evaluated in terms of how such capacities represent a saving to the husband in domestic costs. The value of a home economics course can best be translated as the amount of income which all women who cook at home sacrifice by not working at some other job during this hour, less the amount it would cost to hire someone to cook this meal. Anything which is not part of the occupational exchange system, then, is simply not treatable in this family of models.

Social Models

The next family of models which we want to treat are those derived from sociology. First we will treat the unchanging social models -- those based on a stable or static social system, here using the word "stable" as it is used in demography, denoting a society with a fixed rate of change, and "static" models those societies with no rate of change. The size of a "static" society remains the same over time.

The simplest social model is the "replacement" model. All that is involved in the replacement model is that roles which are maintained in the society, requiring special training, should continue to be filled. It is assumed to be the task of the educational system to fill these roles. The simplest way to see if the school system is performing its job correctly is to enumerate the jobs which exist in society and to be certain that people are available all the time for each job and in the right number.

Thus, in the "static" situation, observe how many priests there were at some time in the past, how many auto-mechanics, how many dog-catchers. When all the jobs that existed to be filled were listed, we then check to see if any of them are under-filled at the moment. If the jobs are under-filled, then something within the school system would be held to be responsible. Not enough men are prepared in seminaries, therefore the number of priests has decreased. Remember that we first treat a static model. The number of jobs does not change.

When we treat a stable model, there should be little more difficulty. We simply most know the growth rate of the society and, knowing this figure, calculate how many men there must be to fill each position three, or twenty, or thirty years from now.

The losses in this model are again dependent on its assumptions. Although we know the quantity of the school output, and the correctness or incorrectness of that quantity, nothing is known about the quality of the replacements. If the school system fails to produce the correct quantity of a certain role, we are instantly involved in all the complicated ideas about "why this has happened." We do not know:

1. if the rewards provided by the role are inadequate;

2. if motivation towards the role is no longer provided by the school system;

3. if people are no longer being adequately prepared for the competencies required by the role.

Nor is there any information about the cost and efficiency of the process of preparation. In a way we have indirect measures of both, but these are probably unsatisfactory. We have first the measurement about filling roles. Are all the roles being filled effectively? This involves us in a calculation of how the roles should be filled. Clearly in a stable society all roles must increase in quantity of role players at a uniform rate.

Most societies are neither precisely static nor stable. In most stable societies, a s the United States is more or less, there is another complication--that of changing role distribution. In American society the size of various roles -- the number of people who occupy them -- is changing. The role of farmer, for example, is getting smaller and smaller; the role of Roman Catholic priest, while not getting smaller, is growing at a slower rate than it once did.

This makes the role distribution of the society more complex, and makes it more and more difficult to calculate the number of people who should be moving into roles. It is often difficult to calculate just how many people there should be ideally in a given role at a given time, unless we consider the growth trends of that role in the past. When we begin to do this we find that the growth trends are dependent on the number of people who enter the role. This becomes meaningless when we discover that the number of people who should enter a role is entirely dependent for its calculation on the number of people who do enter a role.

This is not always true. The growth rates of some roles are entirely dependent on the relative rate of growth of other roles. Thus, the number of managers which will be needed is dependent on the growth rate of the industrial labor force, or on the growth rate of automation.

But, whenever the role distribution is changing, or whenever there is no other factor which determines at least in part the desired size of a role, we cannot calculate the ideal size of the role; simply because our figures are dependent on, and therefore no different from, its actual size.

Thus we will usually be able to predict the size of a role at some time in the future. But we will not know if this is the ideal size of the role or not. We will not know if there is merely a mistake on the part of those responsible for filling the role, or if there is a changing role distribution.

Similarly in calculating cost, there are manifold difficulties. In this method of role-based education, we may find that one of the costs of imposing such an educational system is a high percentage of dropouts from the school system. If the educational system is sufficiently role-directed, and if the roles are undesired or unattainable, then there will be a high attrition rate among those not interested in the roles prepared for.

We could also calculate the economic costs and the economic gains. The cost of education (perhaps as calculated by Machlup, taking into account other kinds of non-school training) would be a simple figure. The productivity or the output of the school system would be another problem; the output could only be calculated by considering the actual distribution of roles, and comparing this somehow with what would be produced under the ideal distribution of roles. (As yet no easy way to do this is available.)

This, of course, gives only the economic costs. It does not consider the costs of replacing such roles as father and mother, which are only indirectly productive -- they are infrastructural roles in a way. If we consider the economic system as the end product of the society, then we can treat such roles as to some extent sine qua non of the occupational system. The interrelationships between family stability (as reflective of adequate role preparation for parenthood) and economic productivity would then need to be studied.

We have, however, no adequate theory of measurement purely in terms of replacement of all roles, because we have no theory of the value of such roles. Simply stated the reason is -- we could interpret the value of social roles in societal terms, if the society were clear in its terms of social value. Even if we did this however, we would have no idea, whether the methods the society chose for the attainment of its own values were the best methods. Only if the society were to say, as many have in the past, that the production of "x" is good in itself, could we be certain that the society were actually doing well.

We are hampered in these discussions by the problem of hierarchically arranged values, i. e. some things exist for the good of other things. This exposes us to questions of how efficiently we are producing the "other things."

Social analysts have posited social integration as some greater good. Those factors which contribute to the integrity of a society are good, those which weaken or destroy this integrity are bad. However, the question of just what this integrity demands are unanswered. Whether they demand stable, unchanging role distribution or a certain trend in role distribution; whether they demand equality of access to roles or equality of rewards for all otherwise unequal roles; whether they demand uniform values throughout or a commitment to a uniform process -- these factors are not decided.

This factor suffers from its inability to discuss value to the individual--the meaningfulness or importance of roles to role players is not discussed. The rewards which the individual grants to himself, the reasons for his selection of roles are unknown and unevaluable.

Social Change

Another family of models are those which see social change as the primary function of education. These models have been most applied to situations in developing societies but we have already seen that some of the material which was originally developed to apply to these societies has application in American society as well. Education, for example, has long been regarded as the pathway to social mobility in the United States. This social mobility has also been seen as a process of social change, but a regularized one, in that the direction and rate of change are usually felt to be predictable. Education then helps some people to change their class status. The direction of change is felt to be towards an equalization of class origins of those in the most highly rewarded occupations. The increase in number of different religious, cultural, and ethnic or racial groups among top-rewarded occupations is the indication of success of this method.

It is clear that the social models are almost entirely non-qualitative. They deal either with the replacement of roles, or with the movement of new groups of people into roles. Therefore, they tell nothing about the individual gains or losses in these movements. Nor can they concern themselves exactly with the value of these patterns. They cannot say whether one role composition is preferable to annother, nor can they consider the benefits of equal opportunity. They must accept the society as a given good, and talk about its maintenance, or they must consider equal access to roles a good and move towards it.

There is absolutely no individual output in such a system. Not only are we saying that individual satisfaction or dissatisfaction is not taken into account, but the career of an individual student after graduation is unimportant except as it contributes to the graphed curve of correct replacement, or contributes to greater equality. The measurement techniques are statistical, and hence dependent on mass populations as outputs -- the position of an individual on his output curve is of no meaning.

Consider the case of an over-filled role, for example. If we consider all the people who together over-fill it, we find that none over-fills it more than any other. Even those who would have filled it just right if taken together are no different from those "added on" people who over-fill it. Some way must be found, on average, to change the motivational structure -- a structure which is implied, not explicit within the model, so that all the people in a role may become members of the right amount needed to fill it.

There is then no unit other than a measurement of the society and its role composition. Either this is in balance, slightly out of balance, changing in a known or unknown direction, but the sole unit is of this magnitude.

Psychological Models

When we turn from additive models to individual models, we gain some clarity and some inclusiveness, but we lose most of the additive ability, and much of the measurability of the earlier models.

There are some psychological models which we could use. They are not specifically psychological in that they deal with accumulated mental health, or anything of that nature, but they are psychological in that they center on individuals and their self-evaluations as the prime output of education. The full range of things which inhere in the individual, which he values, are felt to be the desired outputs of the educational system. The amount of value which the individual places on these products is greater or less depending on the success of the school system in shaping him.

This allows us to take into account all the aspects of an individual's life -- his job, his marriage, his children; the satisfaction he derives from cultural activity; the joy he takes in art. But even as it does so, it forces us to postulate some uniformly measurable single output -- call it satisfaction.

Thus the individual's entire life can be seen as the sum of his satisfactions -- he is married because sex is satisfying; he has children because they are satisfying; art is satisfying. They could, of course, be satisfying in the same way, in terms of our theory. If they are not all satisfying in the same way, it is necessary to return to barter-economy models -- how much sex equals how much art equals how many children? (all to be expressed in some common unit: the satisfaction in "x" spoonfuls of oatmeal per morning -- thus the composite output can be expressed in terms of Picasso-gazing mandays, oatmeal man-mornings, unless some more orgiastic unit is desired).

We must now ask who is qualified to evaluate the individual's total satisfaction, his "happiness"? Does the individual on a scale from one to ten indicate his happiness, or the level of his satisfaction? Or do we profess to have some unit of satisfaction which can be applied in all cases, such as the tendency not to commit suicide, for example.

What about questions of preference and the hierarchy of satisfactions. From the societal standpoint not all sources of individual satisfaction are desirable or even functional. If everyone found his satisfaction in murder, happiness would get higher as the population decreased. Thus happiness must be at least legal. Secondly if everyone found satisfaction in going to movies rather than sitting home and reading books, books and reading ability would very quickly disappear.

There is also a problem about the sources of satisfaction -- people can be taught things, and they can be taught to enjoy things. Thus to a certain extent the structuring of satisfaction is a problem for the school system to undertake. Decisions must be made about what kinds of satisfaction are better than others. Decisions must also be made about what things people will not enjoy at all.

As to cost and efficiency evaluations of this kind of output -- satisfaction -- there is none intrinsic within the scheme of analysis itself. We must observe not only the amount of satisfaction which is produced, if that is what it pleases us to do, but also the functionality of that satisfaction and the prospects for changing sources of satisfaction. Bigotry, for example, may be highly satisfactory; it is socially dysfunctional. Yet it may not be socially dysfunctional for those who are bigots, rather it is more likely to be socially dysfunctional for those against whom they are bigoted. It is perhaps unjust under the theory of satisfaction to challenge the satisfactory state of widespread bigotry. There is no dynamic of change within the theory of satisfaction itself.

There must also be some theory of the additivity of satisfaction -- how best to satisfy a large number of people. Is satisfaction additive? Thus is there more satisfaction if two people are satisfied rather than one? Does it exist negatively, i. e., if one person is satisfied and another is dissatisfied, is the amount of satisfaction in the system one-person worth or no-person worth? Is it better for ten men to be very satisfied than it is for fifteen men to be somewhat satisfied? The critique of satisfaction cannot answer these questions which are basic ones to the additivity of satisfaction. Hence we must combine theories of satisfaction with something else, with ideas of social justice or with ideas of institutional change.

Humanist Models

The perspective which does this best is the perspective of humanism. This is by far the most nebulous kind of theory. Indeed it is often seen as an atheoretical view of man -- not so much in its desire to avoid abstraction, but in its dislike of the implications of modern theory -- scientised empiricism.

There are three kinds of humanistic traditions relevant to education: religious, cultural, and citizenship. These are not really different kinds of humanism; they are simply different kinds of humanistic world views. with some relevance to education. The first may seem the most unlikely since humanism is generally associated with areligious or antireligious tendencies. The sort of humanism here described is one which places its focus on man in the universe -- it emphasizes the universality of human nature, the need for reverence in man, the questing being.

Cultural humanism concerns a "man in Western society focus" -- it emphasizes the continuity of Western society from the time of Homer, through English history and literature, to the present. It is liable to place more emphasis on "the great" than is any other tradition, or at least it places this emphasis more strongly.

Citizenship is a sort of contemporized Vergillian or Horatian attitude. Its primary emphasis is on civil libertarianism, the duties and responsibilities of an intelligent man in his own society. It has a certain element, no doubt, of noblesse oblige.

The strong point of all humanistic approaches is their inclusiveness. They deal with individuals complexly, first as human beings, as human beings living with others, and related to them in a variety of more or less specified ways. Thus there is no conflict in the humanistic tradition between homo sapiens and homo faber. It allows the school curriculum to include most aspects of training and education. It also provides some view of the ideal product of the school system -- depending on the particular tradition involved.

Humanistic approaches then provide:

1. some theory of desired output not dependent on other theories but as an end in itself;

2. a theory of output that is purposely complex, dealing with as many aspects of man's life and his societal life as possible;

3. an ethic of growth, self-motivation, self-selection -- in short, an ethic of self-maximalization, embodying individual choice;

4. thus the theory permits stratified output, but does not predetermine the order of stratification. It allows for equality of opportunity and equality of access to education but does not promise equality or uni-functionality of output.

Humanistic theories of education can be criticized from a variety of viewpoints. Primarily because they are dangerous, not only because "he thinks too much," but because they are often unpredictive. They provide much ground for individual choice. They can be dangerous, too, in that misapplied they are somewhat contradictory. How can one say that individual choice is of prime importance, when individuality is taught -- how can schools continue to emphasize math and science, when the purposes are avowedly directional? The problem can be resolved by saying that even though occupations are highly rewarded by the society, and highly necessary to the society, they may be individually rewarding as well; some experience with them is necessary to decide whether they are individually rewarding. Thus the argument can be reversed -- as long as a variety of options are presented by the school system and no formal pressure exists to force students to take certain courses only, then there is no real argument against including heavy concentrations of societally desired training among them. It is of course pointless and anti-humanistic to provide courses for only segment of the school population and require all students to take them to such an extent that those not interested in a particular field of concentration have no recourse but to leave school. Thus the humanistic tradition argues for the comprehensive high school.

The problem of measurement is still an acute one, however, perhaps more in this tradition than in any other. It seems best to argue from the viewpoint of generalizability.

In judging an individual student one might say -- if everyone were like him, what sort of society would this be? In judging the whole output of a school as a microcosm, one could look at such things as the distribution of occupational preferences -- as self-expressed or as expressed by such things as the Kuder preference tests. If the graduates are strongly unbalanced in one direction there is probably something wrong with the quality of training in other areas. If no one wants to be an English teacher, or a mathematician, if no one even wants to go to college, then the school is not properly performing its function. Occupational programs should evaluate themselves -- what percentage of the occupationally-directed students find jobs upon graduation which are similar to the ones for which they had prepared? What portion of those who worked up to their perceived capacity in the school got the kind of jobs they and the school expected them to receive? If the school compares its distribution of output to a national distribution, and considers the kinds of outputs it has had in the past -- it should come to some sensible conclusions about what its students are doing well and what poorly.

Perhaps, too, the national assessment program, if its aims are sufficiently modest, can be of some use. It should present information about what other students are able to do, and may serve to raise the expectations of high schools, particularly about their own students.

In some remote high schools, much of their poor output could be improved not by higher standards but by higher expectations. Schools which simply set higher standards do not necessarily expect people to meet their standards, they expect perhaps higher percentages of failures and are prepared to accept them. Schools which raise their expectations on the basis of information such as might be available in the national assessment, would believe not only that output should be higher but that output could and can be higher.

Schools, too, should be made to realize that no single track uniform curriculum is going to satisfy everyone, or challenge everyone.

In this way the national assessment program can function as a form of intellectual politic, in very much the same way as the college board entrance examinations. It will serve not so much as a measurement of school output, but as an instrument for changing school output through the dissemination of information. If schools know that the national average on some form of test is 400 and that their graduates consistently score in the upper 200's, there is every reason to see that this can be used as a political instrument from within the school or community for improvement. This is a very different use than has been seen for the test in the past; it is not the sort of instrument which we had wanted, but it is an instrument with a highly important function. It makes highly comparable data (within age groups and between schools) available, for the political use of someone with interests in improving the school's curriculum.

Certainly there will be adverse pressures, and there can be severe bastardizations in the use of the test. If, for example, the test is standardized so that roughly the same material is treated year after year, then it will be possible for the curricula of schools to become adapted to the structure of the exam -- just as is the New York State Regents' exam at the moment, or the French baccalaureate. There is also the possibility that the exam will become so unrelated to anything -- remember there is no check on the significance of anything beyond itself, in the test -- that higher test scores will be negatively correlated with success in life. But with some tempering and some time, and with some attempt to show correlations based exclusively on the test with varieties of success in later life, there could be great usefulness to such an instrument.

The test then, simply because it is a comprehensive test, and simply because it will or could be administered to everyone, will have the effect, if it is well done, of providing an assessment of educational product, with as many elements of transferability and as many kinds of correlation as can be established. But it is important to remember that this is dependent not on anything intrinsic within the test -- it does not measure anything in the ways that it must -- it may simply establish itself as a widespread datum which, if correctly managed over enough time, can be turned into an effective lever for the reform of the school system. It will provide useful, indeed necessary, data for use in any debate about school outcome and school policy. It will become simply by its presence and its irrestible attractiveness a political instrument within the school system, it will be unavoidably important -- it provides a profile of school output from a certain viewpoint -- the consistency of that viewpoint makes it central for organizing many aspects of school output. And its existence may force educators to answer a series of questions about the normative acceptability of the results which are obtained.

REFERENCES

1. Educational Testing Service, "Phase I: National Assessment" (Pamphlet), Princeton, New Jersey, November 1965. Submitted to the Exploratory Committee on Assessing the Progress of Education.
2. F. M. Lord, "Item Sampling in Test Theory and in Research Design," E. T. C., RB-65-22, June 1965.
3. Fritz Machlup, The Production and Distribution of Knowledge in the United States, Princeton, New Jersey: Princeton University Press, 1962.
4. Frederick Harbison and Charles A. Myers, Education, Manpower and Economic Growth: Strategies of Human Resource Development, New York: McGraw-Hill Book Company, 1964.

C. CONCLUSIONS AND RECOMMENDATIONS

8. Toward Evaluation: Some Indispensable Elements in the Measurement of the Effectiveness of Educational Systems

Our principal aim throughout this enterprise has been to develop a model for the measurement of the effectiveness of educational systems that would be culturally specific enough to be used in this country, and culturally general enough to be used for cross-national comparisons, albeit with some adaptations. Most generally speaking, we may say that we have failed in this effort, in the sense that we will not terminate this report with a bright, shiny new model.

In another and important sense, however, we may say that our effort has yielded a certain kind of success insofar as we have learned a good deal of value about educational systems, comparative variations in such systems, and problems that must be solved before we can hope to measure their effectiveness.

Perhaps the single most important new guide to measurement of educational systems that we have learned has to do with the intimate connection between the politics of education, on the one hand, and the model of the system and of its measurement, on the other. Something of this connection is hinted at in our earlier insistence that an indispensable first step in the development of a model of measurement was the assessment of the outcomes of the system under scrutiny. We urged this procedure on the grounds that only if the discrepancy between those outcomes and the intended goals is measured and made a matter of concern, does it then become possible to engender interest and support for more extensive measurements of other aspects of the system.

But the connections between politics and measurement are more intimate and intricate than that. For while it is true that one must engender interest and support before one can proceed beyond simple outcome measurement, the reason is not simply that educators with power will otherwise make measurement difficult. Rather, it is unavoidable in educational systems that a number of publics should have at least slightly differing versions of the criteria of its effectiveness. This is in effect to say that legitimacy and the adequacy of the model of the system that one offers cannot be determined by the "internal logic" of the educational system itself. Rather, that legitimacy and adequacy are functions of the diversity of publics and interests that go into any educational system. If teachers, students, parents, principals, boards of education, and the general public of taxpayers are in effect all members of the educational community, and if their varying versions and interests are unavoidable components of the educational system, then any model of that system must take those diversities into account and include them as part of the field of forces that yields the final outcomes.

More than that, the differing definitions of the adequacy of the system that will be held, unavoidably, by each of these publics, constitute legitimate competing versions of the system, all of which, taken together, represent, for the analyst, the educational system in operation. Thus, if teachers stress humanistic goals, while parents stress "mobility" goals, and students emphasize some combination of humanistic, resource, and mobility goals -- all of those are goals of the system. There is no way in which this can be gainsaid. One cannot appeal to some formal definition of education by which to justify one set of goals over the other. For there is no formal definition of education, at any acceptable level of generality, that will exclude the simultaneous participation in that system of the various publics and their divergent interests.

In this regard, the educational system of any society more closely resembles the political system than perhaps any of the other major institutions. But there is no major institution or subsystem of a society that does not experience the situation that numerous actors in differing roles approach the system with divergent expectations, interests, and criteria of adequate functioning. It is simply that the educational system is perhaps more intricately complex in this regard than any system (except, perhaps, the political). This is particularly the case when the system is in a democratic society, and even more so, when democratic governance of the educational system itself is considered at least ideologically desirable by all concerned. The formal legitimacy of the participation of various publics in the affairs of the system establishes the formal legitimacy of their versions of the system. It is in that inescapable way that the politics of education become intimately involved with the formal structure of education, and requires that any model of the system becomes so attuned.

One can put the matter as follows: the educational system of any society can be said to function as well (or as badly) as it is judged to function by all those who are affected by its outcomes and who are considered by that token to have a legitimate interest in those outcomes.

Once this is said, it becomes apparent that the task of measuring the effectiveness of any such system becomes something substantially different from what it would be if there were agreed upon simple measures of effectiveness, issuing from one authoritative source, with differing versions of that effectiveness considered illegitimate or irrelevant. In turn, this implies that any model for the measurement of the effectiveness of an educational system must make it possible for all the diverging interests connected with that system to discover the extent to which their interests are being realized by the operations of that system.

There is here the further implication that once such an inventory of diverse outcomes has been specified and provided to the interested publics, there will unavoidably be a number of political consequences, in the form of debate as to how effective the system has been overall. That is to say, the various publics will add up the various outcomes in terms of their own arithmetics, and are sure to arrive at divergent estimates of the overall effectiveness of the system. Then the politics of educational interaction will operate to determine the temporary resolution of these competing points of view and new sets of "aims" or "goals" of the system are likely to be formulated. In the normal course of events, one could expect these new sets of aims or goals to reflect the interests of those who are powerful enough to persuade or impose upon the other publics their versions of what should now enjoy priority of attention from the educational community.

This new set of priorities may or may not reflect discontent with the outcomes as measured. And they may or may not be different from the priorities which guided the educational system in its previous phase of operation. Indeed, if the actual experience in American school systems is any sample of what is likely to occur, the greatest probability at any one time is that after much debate the new goals are likely to be substantially the same as the old goals, on the simple grounds that the power to shape these goals is likely to remain in the same hands as those who were previously powerful.

This is even more true, of course, in those European systems where the legitimacy of genuinely democratic participation by divergent blocs in the governance of the educational system, and in the setting of educational priorities, is not nearly as accepted nor traditional as it is in the United States. One can account for the much greater volatility of the American system, and the much greater frequency with which it appears to change directions or priorities, simply on the grounds of the greater actual inclusion of diverse publics in the processes of decision-making. The American system in short does provide structures and mechanisms by which dissatisfaction can be registered and change brought about, albeit against considerable inertia and resistance -- at least significantly more so than is true of most European systems of education. Moreover, the national structure of education in the United States, based as it is upon the partial but effective autonomy of more than twenty thousand separate school districts, makes for maximum diversity throughout the nation, as against the relative homogeneity of local systems in European nations where education is ruled far more monolithically from a national center.

There is, of course, no way of saying which form of governance is more efficient. For the American form makes it just as theoretically possible for divergent interests to be represented in educational efforts as does the European form make it possible for the single-center interests to have their aims realized. Herein lies the greatest single perplexity regarding the measurement of the relative effectiveness of different national systems. For surely the effectiveness of a system must be measured as much in terms of the extent to which its desired political forms of governance realize themselves as in terms of the extent to which the substantive content goals of education are realized within that political framework. That is to say, a major aim of American education, whether the participants know it or not, is to make it possible for numerous and diverse local interests throughout the nation to participate in determining how their schools shall be conducted. To the extent that the structure of local autonomy makes this possible, as it does, to that extent the system has operated effectively, however divergent the achievements of any or all of those locally run systems may in fact be from the content goals they professed to desire to achieve.

The importance of this political "end" of education in the minds of the participants in the United States can be seen in the resistance exhibited to any efforts to centralize the mechanisms of control and in the frequency with which the taxpaying public rejects the budgets and the implied substantive aims of education proposed by the administrative and/or teaching staffs. To keep control of the educational system thus seems to be as high priority a goal of education of local communities in the United States as any of the substantive goals themselves.

The emphasis upon the structure of governance is likely to persist so long as the members of the various communities are concerned that their versions of adequate education shall be represented. So, too, it will persist so long as they feel, as so many of them now apparently do feel, that unless they maintain such local control (and provide evidence of their power by intermittent cutting down of budgets), their notions of what education ought to be doing, and at what level of cost, are not likely to be taken.into account by professional educational leaders.

There is a long, hard, but possible way out of this situation in which educational matters become highly political and educational futures are determined as much by political considerations as by intrinsically educational criteria. In a curious way, the effort at National Assessment, which has engendered so much dispute in the American educational community, represents the beginning of such a way. But precisely because it is not fully cognizant of how broad-reaching that assessment must be, and the political reasons for being so broad-reaching, it is likely to encounter great difficulties. For, the program for National Assessment is attempting to assuage political fears felt by various segments of the educational community through reassurances that the Assessment will not be put to invidious uses. But there is more to the political fears than can be assuaged by that reassurance, however important that promise of freedom from invidious comparison may be. For there is a hard core of resistance throughout the educational community that is based on the fear that varying local versions of educational priorities and goals will not find legitimate expression or recognition in the kinds of measurements that will be attempted by the Assessment program. Specifically, for example, it will not be possible for teachers to claim, from the results of the Assessment, that they have been doing well at their tasks. So, too, it will not be possible for teachers of "soft" subjects to have the fullness of their interests and efforts represented in the outcomes of the Assessment. Similarly, it will not be possible for superintendents and boards of education in small, rural districts to claim relative success at their efforts, given their resources. However, much as it is intended to avoid invidious comparison between local school districts, it seems inevitable that such comparisons will be made. In those comparisons, it will not be possible for those districts that aim at a relatively simpler version of education, more attuned to local community needs and interests, to find reassurance that they have done well at their tasks.

Now, to be sure, any program of assessment or evaluation that aims at making it possible for all participants to be reassured that they have done well is by that token unworthy of being put into effect. There is no point to evaluation except to determine both weaknesses and strengths in school operations, so as to have a basis for strengthening programs where that is needed, and to persist where that seems called for by virtue of demonstrated success.

It is imperative, therefore, that any program of evaluation of educational effectiveness should both take into account the diversity of educational interests, in the first instance, and, in the second, provide for the kind of evaluation that will make it possible for concerned publics to take seriously the measures of their strengths and weaknesses and be moved to appropriate action.

This is obviously a most complex and difficult task. But it can be done -- or at least the first steps in that direction can be taken. We can here specify some of the most general requirements of such a model of measurement.

1. It must include an assessment of the outcomes over a range of diverse subjects or curricular goals, such that the "soft" as well as the "hard" subjects are represented with equal importance.

2. It must include indicators of educational outcome that are of particular interest to the divergent publics involved. Here we refer to the four different kinds of educational orientations specified in chapter 7 of this report, namely, resource development, social replacement, mobility and equality, and humanistic. (We shall specify these in detail shortly.)

3. The evaluation must also include some measure, however rough, of the average gain each student has made in each of his efforts. This means that the assessment of outcomes has to be done at several points in time, so that the change from one time to another can be estimated. The domains in which these changes are measured must be the same as those in which the formal outcomes are measured.

4. It must include assessments by the various publics concerned with the education system of their degrees of satisfaction and dissatisfaction with various aspects of the educational structure and operations, and their reasons for their satisfaction or dissatisfaction. Thus, the evaluation must include, among other things, a series of items for students to answer, in which they indicate what they are satisfied and dissatisfied with in their schools, and why. Naturally, too, teachers, parents, administrators and the general public must also be "polled" in the same terms.

5. As a correlate of 4. above, the various publics must be queried for their knowledge of the curriculum and other school operations, so that their reports of satisfaction and dissatisfaction can be understood against the background of what they know about what actually transpires. The intention here is not to discredit the reports of the less-knowledgeable members of the educational community, but to make it possible for participants to say they "don't know" or "can't say" so that the school authorities will know the extent to which their communities are knowledgeable and the extent to which their intentions and actions have been perceived and understood.

6. It must include at least some measure of community input or support, in the form of a per capita expenditure on education or some other indicator of input.

7. It must include at least some measure of school quality such as percent of teachers with advanced training or any other indices of quality on which there is general agreement.

These are seven minimum requirements for any model for the measurement of the effectiveness of a school system, and the rationales for each of these requirements can now be stated.

The measurement of formal outcomes or achievements in a range of both soft as well as hard subjects is needed so that the various areas of strength and weakness of the school can be identified and so that the teachers and students concerned with these various aspects of school functioning can have their concerns represented.

The inclusion of measurements of outcomes relevant to the four different goal-orientations of education is intended to provide information to the varying publics concerned with school operations to determine how well the school is doing by the criteria they prefer.

The assessment of the relative gains of students over a wide range of subjects is required to insure that the school shall be measured by the changes it has induced in the students, or that the students achieved or failed to achieve, rather than being measured alone by where they stand at any given point without knowing how much change or growth that represents. The information about such change makes it possible to identify areas where the emphasis of subsequent school efforts must be brought most importantly to bear. Most important, the assessment of changes makes it possible for the school to be as concerned with the educational development of the slowest as of the quickest child in the school. Change and growth become the focal points of interest, rather than absolute levels of achievement. This will be crucial especially for those schools who do not turn out a high percentage of "stars" but who may do a very creditable job of moving along large numbers of less than distinguished students.

The reporting by the various publics of their degree of satisfaction is included to make it possible for the system as a whole to locate the areas of content and discontent, and to make it impossible for self-evaluating administrators to stamp approval upon the operations of their schools, come what may.

The determination of the knowledgeability of the publics regarding school operations and outcomes is designed to make it possible for the school officials to improve the lines of communication between the school and the community so as to engender more interest in the operations of the school and to make it possible to recruit support for innovations where these are needed.

The measure of community input is designed to set some realistic limits on expectations when the invidious comparisons between schools begin to be made, as they inevitably will. A school that spends $500 per capita cannot be expected, all other things being equal, to match a school that spends $1,000 per capita.

The measure of school-quality is important because it is most widely believed throughout many educational publics that the quality of the school is more important than any other factor in determining the achievement levels of the children and the amount of growth they can experience.

Is Rational Assessment and Action Possible and Desirable?

The seven requirements for any model of measuring the effectiveness of a school system raise serious problems of measurement, reconciliation of goals, integration of information, and cross-cultural utility. A logically prior question relates to the general question of the role of rationality in the decision-making process. The virtue of any model, whatever its characteristics, rests on the assumption that if knowledge is placed at the disposal of all the publics who have an interest in the educational system their choices about the future courses of action will in some sense be "better" than if they had been compelled to choose in ignorance. This proposition seems self-evident and it has guided our efforts throughout this work. Nevertheless, it is possible to identify certain latent intellectual tendencies, if not a substantial body of opinion, that are suspicious of any precise and orderly attempt to specify educational goals, system characteristics, and outcomes and to establish their interrelationships. This skepticism takes the form of a denial that the scientific knowledge - rational action equation is either 1) possible and/or 2) desirable or 3) if both possible and desirable, appropriate for educational planning.

The first position, in turn, depends on one or all of the following assertions:

1. the scientific method is inherently incapable of yielding valid and reliable knowledge;
2. social events can be anticipated but not controlled;
3. human nature is essentially irrational; and
4. conflicts among interest groups are irreconcilable.

The case against social science rests on the awesome gap between its aspirations and achievements. In their encyclopedic propositional inventory Human Behavior, Bernard Berelson and Gary Steiner themselves social scientists of considerable distinction concede that the behavior sciences suffer from "too much precision misplaced on trivial matters, too little respect for crucial fact as against grand theories, too much respect for insights that are commonplace, too much indication and too little proof, too little genuine cumulation of generalizations, too little regard for the learning of the past, far too much jargon."[1] It would not be difficult to extend this indictment by drawing up a detailed list of fairly central issues in the social sciences which remain unsettled. Indeed in the course of this study we have consistently pointed to the unsatisfactory state of educational research. The possibility of a mature behavior is itself only a grand hypothesis. It is a wager that if we proceed according to methods that have some points of resemblance to those that have been successful in the natural sciences these strategies will also yield benefits to us. A wager is, however, no more than action based on hope and faith.

And yet, and yet . . . there are numerous ways to apprehend the world at various levels of rigor and sensibility but the scientific approach, for all of its limitations, is demonstrably superior to any competing orientation in its capacity to solve some intellectual problems. Experience, insight, intuition, and art perform an honorable function in understanding the human condition but they are by their nature private, that is to say, incapable of certification by publicly certified criteria of verification. There is no way to mediate between two strongly held and opposite intuitions except by judging the intensity of the conviction which in practice usually means resorting to charismatic persuasion or the authoritarian exercise of power by one of the protagonists. Science has no monopoly in generating propositions but only some form of experimental logic and quantitative procedure is capable of confirming them. The scientific style is not the only, nor even a very good, approach to discovery and proof; it is merely indispensable.

There is nevertheless a school of thought which concedes the intellectual power of the social sciences but insists that their ability to understand necessarily exceeds their capacity to predict and their skill in both is greater than anyone's control of the future. The doctrine that men are hostages to history usually assumes that social change is mainly influenced by exogenous factors and unanticipated causes which render rational planning futile. There are numerous examples that contradict the rhetoric of helplessness. Kenneth Boulding cites a particularly dramatic illustration of how apparently modest accomplishments -- internal developments in the science of economics leading to greater theoretical power, the availability of extensive information and the development of imaginative concepts -- have had profound repercussions throughout the entire Western world. According to Boulding:

If one were to look for the most important single reason for the striking contrast between the twenty years after the First World War and the twenty years after the Second, in terms of economic development and the avoidance of great depressions, at least in the developed world, I would nominate the development of national income statistics as the most important factor. The whole concept of the gross national product, for instance, was almost unknown in political discourse before the Second World War. It is true also that certain conceptual changes in the theoretical image of the system, due mainly to the powerful insights of Keynes, went hand in hand with the new information system to create an image in the mindsof economic policy-makers of a controlled market economy, which means that the Second World War represents a real 'system break' in the economic system of the Western world, with a very profound shift in its fundamental patterns of behavior.[2]

The development of a gross educationalproduct may not now be feasible, and its benefits might be problematic but the fact remains that an intellectual orientation capable of transforming economic thought and behavior is in principle capable of developing ingenious devices that can perform the same function for education.

Another version of the doctrine that rational planning is fantasy does not emphasize the inadequacies and impotence of existing scholarship but rather the recalcitrance of the potential consumer. This view stresses the irrational element in human behavior. Of late the conception of man that views him as the captive of his own impulses, that can always detect rationalization but never reason has been buttressed by misinterpretations of recent findings in behavior genetics and ethology. It is, however, the product of a more pervasive tradition.

Thirty years ago Max Lerner observed that "the intellectual revolution of the twentieth century is likely to prove the charting of the terra incognita of the irrational and the extraction of its implications for every area of human thought."[3] It is the world view that emerges in such novelists as Stendahl, Dostoyevsky, and Lawrence and is expressed by such thinkers as Bergson, Pareto, Sorel, and of course Freud. Lerner who is impressed by "the role of instinctual drives as well as of logical formations"[4] nevertheless suggests an effective antidote:

> But does this mean a surrender on our part to the force of the irrational? By no means. There is an enormous difference between the recognition of the role of the irrational and the glorification of it . . . The

work of Freud himself is from this point of view revealing . . . Freud approaches it by rigidly scientific and rational procedures of study. Another example is Thomas Mann, whose novels, such as The Magic Mountain and the Joseph series, are profound explorations of depth-psychology, yet organize these perceptions of the demonic in the human psyche into a framework of values that looks to the life of reason. We have here the foreshadowing of the task of political science in our time--in fact, of the principal task of our age: that of finding a resolution between the necessary role of the irrational and the demands of social rationality.[5]

The final general argument mustered against the utility of ambitious efforts to assess the effectiveness of collective social action does not necessarily assume that individual men are possessed by demons but rather that groups exhibit rather too much short-run rationality. The contention is that no compilation of facts, no clarification of alternatives, no appeal to reason can deter diverse publics from pursuing their own "immediate" and "narrow" objectives. New social inventions such as more precise indicators of educational effectiveness will change the rhetoric of group conflict but not its substance. There is presumably no amount of carefully marshalled evidence demonstrating the dysfunctions of inequality of educational opportunity that could persuade parents in the affluent suburbs to relinquish many of the advantages accruing to their own children for the sake of a transcendent social ideal. Society, then, is ruled not by enlightened self-interest to which philosophical analysis and scientific methodology could make a contribution, but by class and group selfishness in pursuit of transient gain.

This is the message of ultimate despair for if social conflict, the source of social change, is necessarily predatory then the continued survival of stable democracy is an untenable illusion. The capacity of American society to adjust to new paradigms of social reality as increased knowledge and the pressure of events revealed the inadequacy of past formulations--witness the transition from nineteenth century liberalism to the twentieth century welfare state -- testifies to the resilience and ultimate discipline of our political processes. If the United States is not ruled by an invisible hand neither do its people lack the wit to avoid the war of all against all.

Not all those who believe that rational decision-making is possible believe that reason is a desirable guide to social policy. There is a prevalent mood among alienated youth, black militants, and disaffected intellectuals that is antagonistic to reflection, prudence, calculation and exalts feeling, spontaneity, and improvisation. To the extent that this assault on reason finds expression in a formal philosophy it is most closely aligned with existentialism but it is primarily moved by disgust at what it considers the constrictued, "'up-tight' nine-to-five IBM world of the middle-class suburbanite." The method of resistance is to feel deeply, to "turn on," to "do your own thing," to create an art that is but does not mean, to give free expression to emotion, to abjure the "hang-up" of sicklied thought. We need not dwell too long on this syndrome; secular mysticism may provide a necessary corrective for a mechanized society but it is incapable of organizing it for humane purposes.

The final form of opposition to the rational assessment of the relationship between inputs and outcomes comes from those who wish to protect the educational system against its potential traducers. This concern is prompted by the recognition that reliable information released to all relevant publics can result in action that restricts the autonomy of the school and threatens the professional status of the schoolman. It was, after all, the widely circulated reports of differential academic achievement of black and white children that converted community control from a political slogan to a credible educational demand.

At its most responsible the desire to maintain the school as an impregnable citadel where professional judgment is supreme is informed by the conviction that as part of their socialization into their occupational roles teachers learn to care deeply about children and in the most profound sense know better what is "good" for their welfare. At the same time every school system is confronted by community pressures, lay manifestoes on course content, constant scrutiny by newspapers, high level scoldings by college professors and, in general, calls for virtue from all manner of instant expert. The case for the autonomous school rests in the asssumption that professional competence is not a myth and that the community is best served when genuine, rather than counterfeit, experts preside over the schools.

Schoolmen, according to this perspective, are and of right ought to be a priesthood properly jealous of their prerogatives and zealous to guard their secrets. Knowledge shared is ultimately power shared and the nation's children will be the ultimate victims.

This thesis is not devoid of plausibility. We would hazard the guess that in most conflicts between professional educators and their opponents the former have more often stood for a more expert, enlightened, complex, and humane conception of the educational task. But democratic theory is intolerant of monolithic control over areas that are the concern of all and, in any event, massive discontent has converted the school from a professional resource into a beleaguered fortress.

The consolation for those who deplore the politicization of the educational process is that community-school conflict need not be a zero-sum game. Neither loses when the child learns. The problem for research and for policy is to discover the particular spheres in which each can properly exercise sovereignty, and to create the circumstances under which this recognition would be freely rather than grudgingly extended.

The import of this section has been that it is academic in the most pejorative sense to deny possibility and desirability, in principle, of releasing to all concerned publics a reasonably precise account of the degree of consonance between educational aims and accomplishments. The problem is not whether we should try to devise the necessary tools of analysis but how.

The first step may be to develop general decision models which are relevant but not specifically applicable to educational systems. Thompson and Tuden,[6] for example, have developed ideal-typical constructs which may be helpful in defining alternative decision strategies and structures under varying conditions of cognitive and value consensus existing between two social actors or groups. Its major components are 1) *main elements of decisions*, a) causation-consequences of alternative courses of action retroactive and future, and b) preferences about outcomes -- some scale of desirability; and 2) *the presence or absence of consensus among actors or groups*, a) agreement, and b) disagreement. On this basis it is possible to construct a four-field table specifying the emergent strategies and structures.

Beliefs About Causation	Preferences About Possible Outcomes			
	Agreement		Disagreement	
Agreement	Strategy:	Computation	Strategy:	Compromise
	I		III	
	Structure:	Bureaucratic	Structure:	Representative Body
Disagreement	Strategy:	Judgment	Strategy:	Inspiration
	II		IV	
	Structure:	Collegium	Structure:	Anomic

It is evident that only in Cell I, characterized by agreement on both causation and preferences, is the appropriate organization bureaucratic in the Weberian sense. Here the problems are technical requiring only straightforward analysis, i.e. "computation." The rules governing such a structure are in fact the rules of the ideal-typical bureaucracy and include the requirement that 1) specialists may not make decisions outside their sphere of competence; 2) each specialist is bound to the organization's preference scale; 3) all pertinent information must be routed to each specialist; and 4) every issue must be routed to the appropriate specialist.

But as we examine additional combinations of beliefs about causation and preferences it becomes clear that each requires its own specific strategy and organizational structure. Briefly stated the alternatives are as follows:

Cell II - Disagreement on Causation, Agreement on Preferences

Lacking acceptable "proof" of merits of alternatives the organization must rely on judgment.

Issue to be decided by a collegium, a self-governing voluntary group with authority vested in the members.

Rules:

1. fidelity to the group's preference hierarchy.
2. require all members to participate in each decision.
3. route pertinent information about causation to each member.
4. give each member equal influence over the final choice.
5. designate as final choice the rule of the majority.

Cell III - Agreement on Causation, Disagreement on Preferences

An organization facing this problem may fall apart. Compromise is only solution.

All factions should be represented in decision.

Rules:

1. require that each faction hold as its top priority preference the desire to reach agreement, i. e. to continue the association.
2. ensure that each faction be represented.
3. give each facton veto power.
4. give each faction all pertinent information bout causation.

Cell IV- Disagreement on Causation, Disagreement of Preferences

Only hope is the emergence of a charismatic leader.

Must make inspired choice free from conventional restraints. At the same time wish to retain some structure.

Rules:

1. individuals or groups must be interdependent and have some incentive for collective problem-solving.
2. a multiplicity of preference scales and therefore of factions with each faction of approximately equal strength.
3. more information must be introduced than can be processed.
4. each member must nevertheless have access to major communications networks in the event that inspiration strikes.

Preliminary models of this sort when suitably refined and elaborated can help to define the choices and modus operandi open to diverse groups on the basis of their goals and their estimates of the consequences of alternative courses of action. Their decisions will be more rational in the educational sphere if they have at their disposal a model for the measurement of the effectiveness of a school system whose major components we have previously identified as outcomes, reactions by diverse publics, average gain per student, community input, and indices of school quality. As we shall see such a model is not free of difficulties, but none that cannot be overcome.

How Shall All These Things Be Measured?

When one suggests such a diverse array of things to be assessed or estimated, the natural question that arises concerns the available technology for the required measurements. It is quite obvious that aside from the rather well established methods of instruments for measuring certain limited kinds of cognitive achievement, the measurement of school performance is in a relatively primitive state of development, and the measurement of such things as school quality is perhaps even less well developed.

But in principle these undeveloped domains can be developed and without excessive difficulty, or at least without any greater difficulty than was experienced in developing measures of cognitive performance. What is required is that educational specialists be willing to make as many simplifying assumptions about the boundary lines of the "domain" under consideration, and about the indicators they will suggest for these new domains of educational outcome.

Surely it is true that we are only at the beginning stages of measuring such features of human functioning as "creativity." But almost all would agree that some measure of creativity must be developed, and that we can no longer tolerate that kind of resistance to measurement that arises out of the simple-minded plea that such things in principle cannot be measured. If we use the term, as we do, and if we think it stands for something important, as we do, and if we presume we know what we mean by the term when we use it, as we do, then all the major ingredients for beginning to develop measurements of the phenomena are in hand. The same remarks must be made, without alteration, for such other phenomena as school quality, achievement in art, and other soft subjects, and changes or growth patterns in these domains of educational experience.

Whatever the pressures may be to "get in to the field" with some form of educational assessment now or tomorrow, the fact remains that sometime, and preferably sooner rather than later, we shall have to develop such measures if we are going to speak with any sense of measuring the effectiveness of an educational system. If we must measure effectiveness in some domains before we have developed all the necessary instruments, let us be frank to admit our inadequacies and to indicate our intention to eliminate them at the earliest moment possible.

How Shall All These Diverse Measures Be Added Up?

The answer to this question is that they shall not be added up. Rather, we think in terms of developing profiles of accomplishment of any given school on which are represented their achievements on the various dimensions indicated above. Nor will any claim be made that, for instance, one unit of formal curricular achievement is equivalent to two units of achievement in growth or vice versa. These are not to be considered additive elements. They are to be seen rather as separate but related elements in school performance, which together give a picture of a school in operation. It will then be up to each public, or each individual who scans the profile of evaluation to determine for himself, by his own values, and his own preferences for one as against another type of outcome, how well the school is doing.

Must the Measurements All Be of the Same Kind?

The answer to this is "no." The measurements do not have to be of the same kind, nor enjoy the same quotients of reliability and validity, nor be calculated along the same kinds of dimensions of scales or continua. Indeed, one would expect that the ways in which achievement in the art and music classes will be estimated will be significantly different from the ways in which the achievements in mathematics or English will be measured. One need not here and now suggest the specific ways in which measurement in these soft subjects will be accomplished. But it is evident and salutary that in the process of arriving at some decision as to how to measure educational development in these fields, the teachers and specialists in these fields will have to be deeply involved in the decisions so that their aims and goals in their aspects of education will be reflected in the measurements.

There will, of course, be serious and severe disagreements and conflicts while such issues are being debated and worked out. That is only to be expected. One can think, for example, of the decisions that will have to be made regarding such issues as whether the development of children in art should be measured by the artistic quality of their products, or the psychological gratifications they evince, or the growth in personality they exhibit, or the slide-over and carry-over effects into cognitive domains they may begin to manifest after exposure to the art curriculum. Any and all of these are legitimate claimants on some degree of attention as to their possible utility as measures of outcome of art education. It may be that all four dimensions of the art experience will be included in the evaluation, and perhaps even others. But the fact of such diverse possibilities should not be taken as too difficult to overcome -- not, that is, if one wants to begin to develop reasonably sensible ways of measuring educational effectiveness.

What Shall Be Done About the Diverse Goal Orientation?

Here reference is made to the differences among those educational orientations identified as human resources versus social replacement versus equality-mobility versus humanistic. In an earlier chapter the essential differences among these orientations were specified. But these may be summarized once again at this point.

Briefly, the resource-development model is primarily concerned with the school as an agency for identifying, recruiting, motivating, and training highly specialized talents in a variety of subjects, most particularly those deemed most relevant to overall social needs of the moment, e. g. skilled methematicians, or specialized musicians, or whatever. Every society is concerned more with one kind of talent than another at various points in its history. Especially in a democratic society, the society tends to turn to its public school system for the creating of these talents. It asks the schools to identify and train the elites it most urgently needs. Every school system should therefore be able to estimate at any given point how well it is responding to this need for specialized, elite talent.

If this insistence seems gratuitous in view of the fact that most schools appear to be concerned almost alone with such elite talents now, it must be answered that while this may be true, that concern for specialized talents is not likely to diminish at any point and every school is likely to be answerable to some portion of its constituency or its wider referen'ial publics for these talents. Nor can it say that it does not wish to respond to these demands upon it for the training of such specialized talents. It may wish to provide more balance than is today normally practiced between these demands and those for the education of all of its students to a greater degree of equality. But that is a matter to be politically articulated elsewhere, and does not obviate the likelihood nor the responsibility of answering this demand.

By contrast, the social-replacement orientation of education is concerned with the general division of labor and the supply of talents for all echelons and levels of that division of labor, rather than with the specialized and more rare talents that form the focus of concern of the resource development model.

Here what is being asked of the school is that it concern itself with the replacement of the society or with refurbishing and replenishing the supply of all kinds of labor. This is a very different kind of demand than the demand for specialized talents, where the quality of the education is measured by what it does for its talented elite and how many such talented elite graduate. Here, in the social replacement model, the quality is a function of how much the school does for all its students, by way of preparing them for later adult roles, and, in turn, how much it does for the society at large, by preparing citizens at all levels of capacity to take up their places in society.

The third model -- that of equality and mobility -- is concerned primarily with the extent to which the school is making it possible for children from a variety of social class backgrounds (or ethnic, or whatever) to learn, and how much of the relative disadvantages that certain children bring into the schools with them are being overcome by virtue of the school's efforts, and how well, finally, the school is making it possible--within the limits of its possibilities--for children to take up places in the adult division of labor which differ from those held by their fathers.

The fourth model -- the humanistic orientation -- is concerned with the "whole" child: his affective as well as cognitive being, his sensitivity to the arts as well as his capacity for higher mathematics, his capacity to live easily and beneficially with people who are different from him in background and talent as well as his ability to perform specialized occupational or educational roles. The concern here, too, is with the intrinsic satisfaction the child experiences from his education. Are his school experiences gratifying in themselves without reference to any instrumental value they may have? Is school fun for the child? Does he experience a sense of personal growth and development out of his activities at school? All these and other related questions are relevant to the humanist approach to education.

As can be seen, each of these four approaches stands for a different conception of the educational process and what one or another educational public thinks the schools should be aiming to achieve. Normally, these are the models which are debated as though they were mutually exclusive. To some degree, they have been mutually exclusive in the sense that when a school opts for one rather than the other, it tends to produce one rather than the other. But it is apparent that it is possible for the school to be simultaneously concerned with all four kinds of outcomes, and that, under specifiable circumstances, the experiences relevant to all four sets of outcomes could be so organized in the structure and functioning of the school as to make them reciprocally and mutually supportive of each other. They are antagonistic, we suggest, only to the extent that they are viewed as such, and to the extent that schools deliberately set themselves to achieve one rather than another.

In sum, all four are to some extent realizable without serious interference with each other and indeed with possible mutual facilitation; all four represent "reasonable" demands that the society or more immediate educational publics may in good conscience make upon the school; all four can be shown to be of value to any democratic society; the focus of any school upon all four is not likely to require significantly larger sets of resources than the exclusive concern on just one or two of these outcomes.

How Are the Achievements of These Diverse Goals To Be Measured?

The answer to this most difficult question once again involves the requirement that educational measurement specialists be willing to make some simplifying assumptions regarding the domains that these four orientations signify and the indicators they are willing to use in assessing achievement within the domain. Most of the requirements for the measurements here have already been specified in general terms when the necessary ingredients of any educational model were earlier specified. Particularly, we refer to the requirements for measurement of objective achievements and changes from one time to another across a range of subjects and experiences; and to the subjective reporting by students and teachers of their degrees and qualities of satisfactions with the educational experiences they are having.

Little mention has been made, however, of possible indicators of "equality and mobility" achievements of the schools. Here, however, one can satisfactorily utilize some aspects of the measurements of changes. For it is in the equality of effects, as measured by changes, that equality of educational impact is truly to be measured, rather than in equality of actual level achieved. Given even the condition in which only differences in natural talents separated students from each other, we should still be concerned with the equality of their educational experiences and this could be measured only with some reference to how much chance did each one have to grow, given his different starting point.

"Mobility" measures would involve certain additional readings the schools would have to take, on such matters, for instance, as changes in occupational aspirations over time, among the students, changes in college or non-college plans, changes in percentage of curriculum devoted to college-relevant versus college-nonrelevant subjects. So, too, shifts in self-image could conceivably be most important here especially those which would testify to the extent to which the child has grown from one who, by various indicators, had shown himself earlier to feel incapable of mastering or even attempting certain educational tasks, whereas by comparison he now sees himself as one who is quite up to trying these new ventures, with some spirit of confidence in possible success at them.

This is not the time or place to suggest in any further detail the kind of possible indicators one might develop for each of the four major orientations into which all educational goals can be classified. Suffice it to say that some reference to each of them seems indispensable in any measurement of educational effectiveness, if one is to observe the right of diverse sectors of the educational community to secure information regarding the outcomes in which they are most interested.

How Can All This Information Be Integrated?

The question raised here refers to the possibility that some analyses might be made of the relationships among these diverse outcomes such that one could come increasingly to understand not only what was being accomplished by the schools but what factors were responsible for its achievements or lack of them. The specifications of the required elements in such measurement, earlier stated, include at least a minimum number of such analytic elements. Thus it should be possible to analyze the relationship between various input factors -- community and student -- and school outcomes, and between various kinds of school outcomes themselves, and between certain measures of the quality of the school and school outcomes.

At the outset of any such measurement effort, however sophisticated the integrative analyses that are attempted, only the most simple versions of the connections among these various aspects of educational enterprises should be communicated. The purposes of such analyses and communications are several. They include informing the schools themselves as to "how well they are doing" and, in part, why and how it is they are doing that well or poorly. Such analyses also make it possible for the communities in which schools are located to become both more knowledgeable about the how and why of the achievements of their schools and the extent to which and ways in which they as communities might more effectively participate in improving the quality of education that is being provided. Finally, such analyses make it possible in the long run for research into educational processes and outcomes to become a standard feature of education itself. We refer here to the abiding need for continuous evaluative research which serves both basic and applied purposes. If we are to develop models that more adequately and fully represent the domains of educational achievement in which we are interested, this can only come about when the value of such evaluative research comes to commend itself to the concerned publics, such that they come to recognize that domains of education other than those which they have heretofore been narrowly focused on are in fact of some real significance and value.

The possibility of developing alert educational communities, consciously and rationally concerned with the relationships among input, processes, and outputs in their educational enterprises, depends importantly on that kind of development of awareness and knowledge. For it is evident that the diverse interests that these various publics now represent are likely to be matters of continuing debate, and claims of mutually exclusive value, with concomittant denial of value to other orientations, so long as they represent "either - or's" of educational policy. It is only when they are seen as separate but connected aspects of the same enterprise that the conflict approach to diverse educational orientations has a chance of being replaced by an approach that recognizes the importance for all of some significant attention to the large range of possibilities that effective education can serve.

The Cross-Cultural Utility of This "Model"

The components that were earlier specified as minimal requirements for any model of measurement of school effectiveness seem to us to be relevant, without qualification, for virtually any type of school system within the Western European culture pattern -- at least insofar as we have been able to determine. The variations in school systems in this culture-area are such that the greatest difficulty in applying this "model" will arise from unequal concern with one or another of its aspects. Thus the more traditional systems are likely to be not nearly as concerned with student reports of their subjective feelings of satisfaction. Yet political events in each of the societies are likely to make them aware of the need to be more concerned with student reactions than they have been in the past.

In the same vein, various school systems may be more concerned with resource development or social replacement than with equality-mobility and humanistic versions of education. Their lack of concern, indeed, may be such as to cause them to exclude measurements of such aspects of educational process and outcome in their own evaluations. The absence of such measurements can then be taken as indicators of degree of concern and the appropriate conclusions may then be drawn regarding the comparative outcomes on these dimensions of the schools involved.

If international reporting is open enough and sufficiently widely disseminated, it should not be long before the unevenness in the kinds of concerns of the various school systems becomes apparent. At that point, it may become possible for school and government officials mutually to encourage each other toward the development of concern for certain types of outcomes that other schools exhibit that their own do not. It is the hope that by this process of dissemination, exchange, and comparison, the school officials themselves may be "educated" toward more common goal orientations in their educational enterprises.

So, too, in making international comparisons it will be possible, through the use of the model described, for school systems to compare themselves on their relative strengths and weaknesses, such that any given school may discover some area of competence or excellence in which it can take pride, while noting those areas in which there are comparative deficiencies. By providing points of pride to which systems can point, it is hoped that thereby the areas of weakness can more tolerably be perceived and more sympathetically reacted to. Since there is no mechanism for coercion of schools toward more uniform orientations, nor should there be one across national boundaries, this model for the development of awareness and concern is perhaps the most effective that can now be suggested. The same kinds of observations apply to comparisons of schools within one country, of course, insofar as one wishes to refrain from employing coercive sanctions to secure more uniformity and would prefer, instead, to rely on effective agents of education and persuasion.

Finally, we may call attention to the possible utility of including in the measurement of school effectiveness that approach envisioned in the concept of Gross National Educational Product, in which whatever is subject to economic translation is in fact translated, and dollar values on investment and outcome are estimated. The objections to this kind of approach in the past have centered upon the extent to which such measurements in dollar terms violate the basic tenets of that educational philosophy which decry the notion that education should be profitable or that educational values can be given equivalent dollar values. We have seen how, in fact, many educational values cannot be given dollar translation, without significant loss of important educational content. We have seen, too, how the absence of a common market for educational outcomes, and of a common unit for their translation, makes the concept of Gross Educational Product something less than desirable, if and when it is proposed as the only way to get some rigor into educational measurement.

But if the approach to education through GEP is only one of many others that have been recommended here, and if the dollar estimates of input and outcome are seen in the context of other inputs and outcomes that are not translatable into dollar values, and if therefore the qualitative dimensions of educational transactions are thereby enabled to preserve their integrity, there can be little objection to learning as much as possible about the economic aspects of that transaction. Indeed, it is to be expected that at critical points, two programs otherwise equally appealing in their prospects may be able to be compared for their relative costliness, stated alone in economic terms, such that, where no other factors make it possible to choose between the alternatives, the costs of each may become properly determinative of that choice. The danger lies in the greater appeal of the precision of the dollar sums as against the indecisive character of qualitative assessment. Here, however, it is up to school and governmental authorities and any other concerned publics to come to understand why they must not be beguiled by the appeal of precision into discarding their concerns for other, less rigorously measured, outcomes of education. The danger, in short, lies not in the economic measures themselves but in their misuse and in the subsequent abuse of other educational orientations. When such economic measures are used on a cross-national basis, in conjunction with other measures of national educational output, we shall be at a point where we will be able better than ever before to make international comparisons that are meaningful in their content, significant for a wide range of educational orientations and capable, when disseminated and reacted to, of generating some common concerns across national lines for the quality of education in all countries.

References

1. Bernard Berelson and Gary A. Steiner, Human Behavior, New York: Harcourt, Brace, and World, Inc., 1964. p. 11.

2. Kenneth E. Boulding, "Reality Testing and Value Orientation," International Social Science Journal, Vol. 17, 1965, No. 3. p. 410.

3. Max Lerner, "The Discovery of the 'Irrational': Personal and Collective," in Warren G. Bennis, Kenneth D. Benne, and Robert Chin (eds.), The Planning of Change, New York: Holt, Rinehart and Winston, 1964. p. 117.

4. Ibid., p. 119.

5. Ibid., pp. 119-120.

6. See James D. Thompson and Arthur Tuden, "Strategies, Structures, and Processes of Organizational Decision," in James D. Thompson, et al. (eds.), Patterns of Administrative Leadership, Pittsburgh: University of Pittsburgh Press, 1959. pp. 195-216.

D. APPENDICES

1. Aims and Outcomes of Education: A Selected Annotated Bibliography

AIMS AND OUTCOMES OF EDUCATION

(Selected, Annotated Bibliography)

I HISTORY - Background Readings on the Aims and Outcomes of Education

Members of the Faculties of Andover, Exeter, Lawrenceville, Harvard, Princeton and Yale, General Education in School and College: A Committee Report. Cambridge, Massachusetts: Harvard University Press, 1952. 142 pages

The School and College Study of General Education, supported by the Fund for the Advancement of Education, Ford Foundation, was conducted by one faculty member from each of three schools (Andover, Exeter and Lawrenceville) and one from each of three Universities (Harvard, Princeton and Yale). The Committee undertook to define the goals of a general or liberal education and to design a program to reach these goals. Through a study of academic records, surveys of course content and methods of instruction, student questionnaires, and panel discussions, the Committee hoped to demonstrate that the work of the superior student in school and college could be integrated into a continuous process encompassing inquiry into many fields as well as study in depth in areas of concentration. In this way, waste in the form of duplication of study in high school and college could be avoided and lack of motivation caused by boredom could be prevented.

A detailed curriculum for grades 11 through 14 and covering the major areas of a general education is outlined. A proposal is made that students, the educational institutions and the society would be benefited if certain highly qualified students were encouraged to complete their education in seven years, either by omitting the last year of high school or by being admitted to college with sophomore standing. An appendix contains a proposal for an advanced placement experiment and several excerpts from the replies to the student questionnaire.

Butler, Nicholas Murray, The Meaning of Education. New York: Charles Scribner's Sons, 1915. 378 pages

In this comprehensive volume, the author deals with a wide range of educational problems ranging from the philosophy of education to discipline and the organization of kindergarden. In the title chapter, he contends that spiritual growth should be the ultimate aim of education. In other chapters, he traces the history of education, discusses principles of education in the United States and reviews organization and methods for all educational levels, from kindergarden through college or university.

Dewey, John, "My Pedagogic Creed" (pamphlet originally published by E. L. Kellogg and Co., 1897) included in Archambault, Reginald D. John Dewey on Education: Selected Writings. New York: The Modern Library, 1964. 439 pages

In this work, John Dewey sets forth many of his innovative ideas on education. He defines education as the fundamental method of social progress and considers the place the school has in the total education of the child. In evaluating subject matter of education, he states that as school is primarily a social institution, the child's social activities must be of central importance in his training. The interdependence of education and society is underscored throughout this as well as the other works in the collection.

Harvard University Committee on the Objectives of a General Education in a Free Society, General Education in a Free Society. Cambridge, Massachusetts: Harvard University Press, 1945. 267 pages

The second World War forced educators to reevaluate the relationship between education and society and the aims and methods of education. James B. Conant, president of Harvard University, charged a committee of faculty members of the arts and sciences of education to examine the usefulness of general education for both high school and college students. In its report, the committee distinguishes between the collective needs of the society and the diverse needs of individuals. Its answer to these dichotomous needs is a basic general program with several levels of difficulty. Implications for high schools and colleges, and for Harvard University in particular, are explored.

Hutchins, Robert Maynard, Education for Freedom. Baton Rouge, Louisiana: Louisiana State University Press, 1946. 105 pages

Dr. Hutchins here expands the Edward Douglass White lectures delivered at Louisiana State University in 1941. He defends his thesis that a liberal education should be obtained by every citizen of a free society. He states that the foremost purpose of education should be to develop students with a "social consciousness and a social conscience." Schools should attempt to produce graduates "who want to improve society and who know how to go about it."
Dr. Hutchins also discusses the University of Chicago plan for a two-year program leading to a Bachelor of Arts degree.

Maritain, Jacques, Education at the Crossroads. New Haven: Yale University Press, 1943 120 pages

Mr. Maritain has based his book on the Terry Lectures, delivered at Yale University in 1943. The first part of the book is devoted to the aims of education, more specifically to the desired characteristics of an educated man. He stresses moral, spiritual and intellectual aspects of education over vocational aspects saying that practical aims are best provided for by "general human capacities." Other lectures explored the dynamics of education, the humanities, liberal education and problems arising from World War II.

Thwing, Charles Franklin, What Education Has the Most Worth? New York: The Macmillan Company, 1924. 225 pages

Dr. Thwing, President Emeritus of Western Reserve University, attacks the concept he believed prevalent at the time that the most desirable form of education was the one centered on natural sciences. The author argues that "education of the intellect" is of higher value and that "human studies" education leading to moral behavior towards others and for the society are at least as important as "science studies." He outlines probable results to individuals and to the society from such an education as he proposes. Sections of the book deal with new technologies of the day, such as moving pictures, and their effects on the society.

Whitehead, Alfred North, The Aims of Education. New York: The Macmillan Company, 1929. 247 pages

In the title article, the author states that education should strive toward activity of the mind and not toward the filling of minds with "inert ideas." He believes man should possess both culture (". . . activity of thought and receptiveness to beauty and humane feeling") and expert knowledge in a chosen area. The remaining essays cover such topics as the rhythm of education, technical education, classics, mathematics, universities, organization of thought, and science.

II AIMS OF EDUCATION

Beezley, P. C., *Education for What?* New York: Bookmailer, 1963.
208 pages

Mr. Beezley is concerned by what he feels is a lack of understanding throughout the teaching profession about the true nature of communism. He delegated SPX Research Associates to conduct a survey of commissioners or superintendents of education in all 50 states as well as state, county and local officials. In this book, the author indicates why he felt the study was necessary, how it was conducted, the results of the survey, and recommendations for action based on the findings of the survey.

Bestor, Arthur, *The Restoration of Learning.* New York: Alfred A Knopf, 1956. 459 pages

Following his criticism of the trends of American education in *Educational Wastelands,* Dr. Bestor here offers his views as to the direction education should take. He contends that education, with its major objective "adjustment for life", has become aimless and that a return to a central emphasis on academic disciplines is in order. He feels teachers are almost uniformly inadequately trained in the disciplines they teach and that overly stringent state pedagogy requirements prevent many potential teachers from entering the field. Proposals to counter undesirable trends and educational problems are offered.

Bloom, Benjamin S., editor, *Taxonomy of Educational Objectives: The Classification of Educational Goals.* Handbook I: Cognitive Domain; Handbook II: The Affective Domain (David R. Krathwohl, Benjamin S. Bloom, Bertram B. Masia). New York: David McKay Company, Inc., 1956. 207 pages; 196 pages

These volumes present a systematic attempt to provide a means to classify educational outcomes in such a way that the order of elements classified represents a relationship among the elements. Two major purposes of the work are to provide a basis for meaningful communication among those concerned with education and to facilitate comparisons of educational programs and achievements. Volume I, The Cognitive Domain, orders aspects of knowledge, ability and skills by the following increasingly complex categories: knowledge, comprehension, application, analysis, synthesis and evaluation.

Volume II, The Affective Domain, orders objectives such as attitudes, values and interests by the categories of receiving, responding, valuing, organization and characterization by a value or a value complex. Examples and sample test items are given for each category. Volume I concludes with a condensed version of the cognitive taxonomy. Volume II concludes with a condensed version of both taxonomies.

Booth, Wayne C., editor, The Knowledge Most Worth Having. Chicago: The University of Chicago Press, 1967. 212 pages

At a liberal arts conference of students and faculty at the University of Chicago in 1966, representatives of science, the humanities and college administration stated the case for liberal education. This book is a collection of papers presented at the conference and used as a basis for discussion. Several authors develop a historical perspective and turn to Plato and his statements on education for justification of a general education. Others find in today's complex industrial society the need for general as well as special education. Chapters are devoted to such topics as education and the contemporary woman and the role of a liberal arts college within a university.

Brim, Orville G., Jr., Sociology and the Field of Education. New York: Russell Sage Foundation, 1958. 93 pages

This book, one of a series of studies on fields of practice of sociologists published by the Russell Sage Foundation, is a review of research conducted by sociologists in education. The author has culled basic studies from the research literature and organized them under topics of educational importance. After discussing significance of the research findings within each topic, the author suggests directions for further research to elaborate and clarify previous studies or to explore areas not yet examined. Among the topics for which research results are presented are the aims of education, the allocation of materials, the functions of education and the roles of the sociologist in education. Research findings concerning educators are examined both in terms of recruitment, advancement, etc. and in terms of roles in the classroom. Findings concerning students are similarly examined.

Bruner, Jerome S., The Process of Education. Cambridge, Massachusetts: Harvard University Press, 1966. 97 pages

The Woods Hole Conference, upon which this book is based, was attended by scientists, educators, psychologists, and historians. Its purpose was to explore the problems of teaching science and provide the first opportunity for scientists and psychologists to meet for discussions on the psychological aspects of learning. Dr. Bruner contends that children are able to comprehend basic concepts of science and the humanities at a much earlier age than has been commonly assumed. He stresses the importance of "structure" and urges that this underlying structure, or the fundamental principles, of subject matter be presented in a form easily grasped by children and that later training build upon this understanding.

Bruner, Jerome S., Toward a Theory of Instruction. Cambridge, Massachusetts: Belknap Press of Harvard University Press, 1966. 176 pages

Dr. Bruner addresses himself to the problems of achieving the goal of self-fulfillment through the process of learning and assisting students' growth toward using their abilities to achieve a full life and to contribute effectively to the society. He discusses the nature of growth, the role of social scientists in education, and presents a theory of instruction based upon what is known about human learning and the circumstances under which learning is best accomplished. He feels that the teaching of time-and-energy-saving skills (language, mathematics, etc.) is a most important function of schools in light of present day rapid change. He discusses experimental courses in mathematics and social studies with which he has been involved. In a chapter on motivation, he discusses the kinds of rewards inherent in learning and the factors which lead to satisfaction in learning. In a chapter on coping and defending, he states that in cases of learning blocks, the child is often finding ways of avoiding the problem confronted rather than finding ways of solving the problem. He concludes with a chapter on the nature and uses of effective curriculum evaluation.

Clark, Burton P., Educating the Expert Society. San Francisco: Chandler Publishing Company, 1962. 301 pages

In his book devoted to the sociology of education, Dr. Clark examines the "social aspects of educational structures and processes, their social causes and consequences." He is primarily concerned with the functions of education in a technical society that demands a skilled labor force as well as a perpetuation of the cultural heritage. In drawing his conclusions, Dr. Clark refers to research reports and definitive major works in sociology, economics, anthropology, education and psychology. The first topics covered relate education to the society and include Education as a Cultural Agent; Education, Occupation and Status; and Education and Minorities, among others. The latter topics, centering on educational systems and processes, include Organization of the School and College; Student Culture in College; Student Culture in High School; and the Educational Process and the Expert Society.

Conant, James B., The American High School Today: A First Report to Interested Citizens. New York: McGraw-Hill Book Company, 1959. 140 pages

Dr. Conant conducted this critical examination of the American high school with financial support from the Carnegie Corporation and organizational support from Educational Testing Service. He and his collaborators visited schools of all sizes in 26 states, in cities, rural areas, small towns and suburbs. The emphasis of the study is on the comprehensive type of high school which provides a general education for all young people in the community, academic background for those planning further education and vocational training for terminal students. Although Dr. Conant urges the adoption of the comprehensive plan for many communities, he indicates that the needs of the particular community must dictate the type of school instituted. After discussing the general characteristics of American education, the author devotes himself to an analysis of the comprehensive high school, the methods used in his survey of schools, and the results of the survey. On the basis of his findings, he presents a list of 21 recommendations for the improvement of American high schools. Schools that are not comprehensive - the small high school, the large city high school, and the suburban high school - are dealt with separately. Appendixes list states and schools visited, criteria for evaluating a comprehensive high school and descriptions of inventories and educational programs discussed in the text.

Conant, James B., Shaping Educational Policy. New York: McGraw-Hill Book Company, 1964. 139 pages

Dr. Conant here presents his views that traditional methods of determining educational policy have been outmoded by a growing national concern for education and by a revolution in educational techniques. In a chapter on policy-making for the public schools, he recommends methods for dealing with several specific problems confronting education today and stresses the growing importance of policy-making at the state level rather than at the local level. Policy-making for higher education is also scrutinized. He analyzes in detail the school and higher education programs of New York and California and urges widespread adoption of policies similar to those of New York for public education and California for higher education. In the final chapter, the author advocates interstate and regional cooperation in building educational policy, leading to "nationwide" policies rather than "national" policies, and cites examples of attempts in this direction.

Cremin, Lawrence A., The Transformation of the School. New York: Alfred A. Knopf, 1961. 387 pages

The first part of this book is devoted to the sources of the impetus toward progressive education and covers the years 1876-1917. The author reviews the traditions of popular education, the impact of industry on education, the response of educators to immigrants, the underprivileged, rural dwellers and former rural dwellers who have migrated to urban centers. He examines the relation between science and education and devotes one chapter to a discussion of the contributions of early education innovators. The second part of the book, covering the years 1917-1957, reviews the rise of progressive education , its strengths, contributions and drawbacks, its advocates and its eventual collapse.

Cubberly, Ellwood, P., Public Education in the United States. New York: Houghton Mifflin Company, 1919. 506 pages

This history of American education traces the development of educational theories and practices and relates them to the development of the society. Changing educational aims are emphasized. He also discusses contemporary educational problems and outlines possible solutions.

[Educational Policies Commission is made up of elected representatives from the National Education Association, the American Association of School Administrators, the Department of Classroom Teachers, the Department of Elementary School Principals, and the Association for Higher Education. Financial support comes mainly from the National Education Association and the American Association of School Administrators. The Commission studies the major contemporary issues in education and issues reports containing recommendations regarding these issues. The reports frequently center on grave national problems and attempt to demonstrate how the schools can counter or help solve the problems. Several reports analyze the function of the school in a democracy; several in recent years concern the disadvantaged student; certain special areas such as athletics mass communications, and the gifted student have also been examined. A few reports with particular influence or significance are discussed below.]

Educational Policies Commission, The Central Purpose of American Education. Washington, D. C.: National Education Association, 1961.

The 1961 statement of the Educational Policies Commission centers on the importance of rationality to individuals and to the society. By emphasizing the development of rational powers, it is felt that individual dignity and equality of opportunity may best be accomplished. The first section of the report discusses the place of education in American society and the use of rational powers. The second presents research needs and possible methods of developing rational powers.

Educational Policies Commission, The Contemporary Challenge to American Education. Washington, D. C.: National Education Association, 1958. 31 pages

The 1958 E. P. C. statement focuses on the re-appraisal of American education that developed in the aftermath of the Soviet Union's rapid growth in technical ability. The crucial educational needs of the United States at that time were deemed to be scientific education and the knowledge, insights and abilities necessary to the preservation of American democracy and peace. The Commission calls for better education for the gifted, improved guidance and teaching. Short range, intermediate range, and long range programs are proposed. Statistics of education in the United States and in the U. S. S. R. are presented in appendixes.

Educational Policies Commission, Contemporary Issues in Elementary Education. Washington, D. C.: National Education Association, 1960. 27 pages

This overview of crucial educational issues is focused on the elementary grades. The importance of the elementary school, the content of elementary education, and the effects of several variables such as health, environment, etc., upon learning are considered. One section of the report is concerned with implications for policy arising from several issues. The importance of good teaching is stressed and conditions leading to high-quality teaching are outlined.

Educational Policies Commission, Education and the Disadvantaged American. Washington, D. C.: National Educational Association, 1962. 39 pages

The meaning and origins of cultural disadvantage are set forth along with an outline of programs for the public school to combat its effects. A section is devoted to public policy and the education of the disadvantaged.

Educational Policies Commission, An Essay on Quality in Public Education. Washington, D. C. : National Education Association, 1959. 31 pages

In its statement on quality in education, the Educational Policies Commission asserts that the best education is "that which does most to enable each student to develop his abilities and to serve society". The Commission further states that two fundamental characteristics of American education are "universality" and "diversity." The required contributions of the school programs, the school staff and the school administration toward quality education and universality and diversity in education are stated. An appendix contains specific recommendations for salary scales, staff size and per-pupil expenditures for professional services.

Educational Policies Commission, Universal Opportunity for Education Beyond the High School. Washington, D. C. : National Education Assoc.

The E. P. C. proposes that the United States provide at least two years of higher education to all high school graduates and sets forth the rationale leading to this proposal. The report also discusses means by which this universal opportunity for college training may be made possible.

Educational Policies Commission, The Purposes of Education in American Democracy. Washington, D. C. : National Education Association, 1938. 157 pages

A discussion of the democratic process is followed by an overview of significant past attempts to define educational objectives. The Commission then lists and discusses in depth the objectives it recommends: the objectives of self-realization, of human relationship, of economic efficiency, and of civic responsibility.

Fine, Benjamin, Our Children Are Cheated. New York: Henry Holt and Company, 1947. 244 pages

The material in this book first appeared as a series of twelve articles in The New York Times. On the basis of materials gathered in interviews with educators and visits to schools in all parts of the United States, the author presents his views on the problems confronting the nation in the field of education. It is his belief that the greatest weakness lies within the area of teaching. He feels teacher-training is inadequate and that the lack of recognition of teaching as an honored profession leads to low teacher morale and poor recruitment. He lists as major problems teacher shortage, high teacher turnover, low teacher salaries, and low teacher morale. In observing the plight of students, he deplores educational inequalities, whether due to social class differences, racial differences, or rural-urban differences. In the final chapter, the author lists 18 action recommendations "necessary to restore the public school system of the United States to a position of leadership." The list covers such areas as financial support, leadership, teacher training, teaching methods, public interest, and equal opportunities.

French, Will and Associates, Behavioral Goals of General Education in High School. New York: Russell Sage Foundation, 1957. 247 pages

This study is an attempt to describe the goals of a general education in terms of desired behavioral outcomes and is based on "The Survey Study of Behavioral Outcomes of General Education in High School" conducted by the Educational Testing Service in cooperation with Russell Sage Foundation and the National Association of Secondary School Principals with advisory aid of various government and education agencies. Eighteen high school and college level educators individually outlined what they felt to be the goals of a general education and their concomitant behavioral outcomes. A committee of advisors reviewed the goals of the consultants in light of public needs. A committee of reviewers, teachers, administrators, etc., then evaluated the lists of desired behavior based on their background of classroom and school experience. From all the lists was drawn a final list of desired behaviors felt to be important by at least three-fourths of those who reviewed them. The executive editor then organized the final list into areas of desired behavioral outcomes (attaining maximum intellectual growth and development, becoming culturally oriented and integrated, maintaining and improving physical and mental health, becoming economically competent), and the directions of growth involved in achieving maturity (growth toward self-realization, growth toward desirable interpersonal relations in small [face-to-face] groups, growth toward effective membership or leadership in large organizations). A major section of the book deals with detailed descriptions of behavioral outcomes and directions of growth for each. Possible applications of the results of the study for parents, schools, teachers, etc., are discussed. A form for use in evaluating general education programs in terms of behavioral outcomes is presented.

Gardner, John W., *Excellence:Can We Be Equal and Excellent, Too?*
New York: Harper & Brothers, Publishers, 1961. 161 pages

The thesis of this book,written when Mr. Gardner was president of Carnegie Corporation, is that standards of performance of many areas of American life are regrettably low and must be raised if the society is to continue to prosper. In opening chapters, he discusses the conflicting principles of hereditary privilege, equalitarianism and competitive performance. He scrutinizes the concepts of talent and individual differences and attempts to reconcile these factors with equality of opportunity. He sums up with chapters on the meaning of "excellence" and individual fulfillment.

Gardner, John W., *Self-Renewal: The Individual and the Innovative Society.*
New York: Harper & Row, Publishers, 1964. 141 pages

In his second book, John Gardner concentrates on the necessity of fostering creativity in all parts of society for he believes the lack of stimulus in this direction will lead to the eventual decay of our social insitutions. He emphasizes the importance of maintaining one's capacity for growth motivation and adaptiveness past middle age. Only if individuals use all their creative capacity throughout their lives will society prosper as the creative society depends on a creative population. The author discusses growth, versatility, obstacles to renewal, methods of renewal, motivation, attitudes and consequences of decay.

Ginzberg, Eli, editor, *The Nation's Children: 2: Development and Education.*
New York: Columbia University Press, 1960. 242 pages

This volume is the second of three contained in *The Nation's Children*, published for the 1960 Golden Anniversary White House Conference on Children and Youth. Articles were contributed by several authors, each representing his special field of interest and competence. The first three articles cover physical and emotional growth while the latter articles cover education. Ralph W. Tyler discusses educational objectives in a democracy; several authors deal with specific problems of scientific education, employment, wasted talent, education in the armed services, and religion. Gardner and Lois Murphy treat the relationship between the child and society and John W. Gardner urges equality of educational opportunity leading to the maximum development of individual potentialities at all levels.

Goodlad, John, editor, Sixty-fifth Yearbook of the National Society for the Study of Education, Part I

"The launching of the first Russian satellite in 1957 gave us our bete noir, and the schools became our sacrificial lamb. The schools could not remain -- or, at least, could not appear to remain -- as they had been. Change, almost for the sake of change, became the order of the day. Change and the spirit of change have swept along in a disorderly way, proposal following proposal on a broken front. Little time and energy have been expended on appraisal, and adequate evaluative criteria appear not to be available. As a nation, we have not defined what we want of our schools and, therefore, have grave difficulty in seeking to determine the worth of any specific proposal or action."

Goodlad, John, editor, Sixty-fifth Yearbook of the National Society for the Study of Education, Part II

The purposes of Part II of the Sixty-fifth Yearbook are to examine closely and evaluate illustrative recent changes in the schools and to explore in depth changes in related fields to which schools might have or should have responded. Each chapter was writien by an expert in the particular area covered. The first section discusses and appraises changes in the role of the teacher, the curriculum guidance, instructional resources, school organization and school buildings. Section two explores new concepts in the social sciences, behavioral sciences, instructional theory and philosophy, and determines their relevance to education. In a critical chapter, Francis S. Chase attempts to relate the changes in the school to the forces at work in the society as presented in the previous chapters. He notes that many of the changes considered relate to technological progress and have little to do with basic social and philosophic issues arising from the transformation of the society. He points out, however, that other changes, such as those in curriculum development and the role of the teacher, relate directly to culture changes. Appendixes record minutes and proceedings of the annual meetings, lists of members and publications of the Society.

Goodman, Paul, Contemporary Mis-Education. New York: Horizon Press, 1964. 189 pages

In this collection of critical articles on the American school, Mr.

Goodman declares that compulsory education as practiced in the United States leads to conformity which, in turn, inhibits creativity and individuality. He discusses ill effects of schooling in the primary grades, high school and college. He criticizes the avowed goals as well as the content and process of education. He states that goals are too frequently related to spheres outside the areas of school competence. Building on the ideas of John Dewey and A. S. Neill, he suggests alternatives to the present situations which would include community store-front schools of 20-25 students of all ages in urban areas, a two-year period of work or purposeful travel between high school and college and abolishment of grading in college.

Hambridge, Gove, New Aims in Education. New York: McGraw Hill Book Company, Inc., 1940. 226 pages

New Aims in Education, written primarily for the college student, emphasizes the "movement for individualism", using a format of letters to a college student. Mr. Hambridge attempts to help the student set educational goals for himself and to find the best ways to reach these goals. He uses the "Pennsylvania Study" of higher education sponsored by the Carnegie Foundation as a basis for his observations and recommendations. Several chapters introduce standardized tests, their uses and a smattering of statistics.

Hanna, Paul R., editor, Education: An Instrument of National Goals. New York: McGraw Hill Book Company, Inc., 1962. 210 pages

The 1961 Cubberly Conference at Stanford University used as its basis the Report of the President's Commission on National Goals. The ten papers presented at the Conference are reprinted here with a preface by I. James Quillen and an introductory chapter by Paul R. Hanna. In his chapter, Hanna discusses two central issues of education: shall education be thought of as consumption or as an investment; and, should education be primarily concerned with "process" (how to think, how to develop rational powers) or with "content." He then outlines the proceedings of the Conference. Frank Pace discussed national goals in relation to world problems. Neil H. Jacob, Claude A. Buss, James R. Killam, Jr. and Max Lerner then covered in more detail the application of national goals to various specific areas of the society; domestic, economic, scientific and humanistic. Three Stanford University educators, Quillen, Robert N. Bush and W. H. Cowley, then suggested ways the national goals could be translated into specific educational objectives for elementary schools, secondary schools, and institutions of higher learning. J. W. Fulbright contributed a paper on National Goals and National Consensus and John H. Fisher gave the final paper which discussed the role of education for all of the nation's goals as outlined by the President's Commission.

Hechinger, Fred M., editor, Pre-School Education Today: New Approaches to Teaching Three-, Four-, and Five-Year Olds. Garden City, New York: Doubleday & Company, Inc., 1966. 150 pages

This book, conceived as a primer for those instituting pre-school programs, covers basic background information on disadvantaged children and various useful educational programs for these children. The book was planned by the Council for Public Schools, Inc., a national group of educators and laymen sponsoring the development of new courses of instruction for public schools. Fred M. Hechinger, education editor of The New York Times edited the book and contributed the first chapter, which deals with the use of education as a means of reaching the goal of equal opportunity for all American children. Three articles focus on the effect of social and psychological factors on learning. Martin Deutsch contributed a paper on early social environment and on facilitating development in the pre-school child. J. McVicker Hunt contributed a paper on the psychological basis of using pre-school enrichment to reduce the effects of cultural deprivation. The remaining articles, written by educators, focus on types of pre-school programs. With the exception of the Hechinger article, the papers have been published previously or have been presented at professional meetings.

Hook, Sidney, "Perennial and Temporal Goals in Education", The Journal of Higher Education, January, 1952, Vol. XXIII, No. 1, pages 1-12

Professor Hook offers a distinction between "descriptive" and "normative" elements of education. Confusion results, he believes, when educators include the educational goals they favor within statements intended merely to describe education. He expresses the desirability of first describing the process of education and then establishing the goals or the direction of education. He distinguishes three overlapping types of educational objectives: powers, skills and techniques; subject matters, fields of study and interest; and personal and moral habits, values and loyalties. He then enumerates perennial goals, those to be attained by all college students, within each of the three categories of objectives. Goals he believes necessary for students in the coming decade or so are similarly treated. He concludes by stating: "The best goals cannot ensure the best education and the best education cannot do everything. The best it can do is to make the individual capable of self-education."

Kearney, Nolan C., Elementary School Objectives. New York: Russell Sage Foundation, 1953. 189 pages

The Mid-Century Committee on Outcomes in Elementary Education, composed of a committee of consultants, a committee of critics, and a survey committee, carried out a study designed to define the goals of elementary education described in terms of desired behavior. The study, of which this book is the final report, was sponsored by Educational Testing Service, Russell Sage Foundation, The United States Office of Education, and The Department of Elementary School Principals of The National Educational Association. After each of the committee consultants, composed of education leaders, described desired behavioral outcomes of elementary education, the committee of critics (teachers and supervisors) evaluated these goals using their knowledge of the classroom situation. The survey committee (representatives of teachers, researchers, administrators and the public) then organized the goal recommendations within a framework of four types of desired behavior change for each of nine curriculum areas. The plan and operation of the study are outlined and each desired behavior change for each curriculum area is described in detail. Possible applications of the study results for schools and the public, for educational theory and for educational measurement are discussed. Types of behavioral change: knowledge and understanding, skill and competence, attitude and interest, action pattern. Curriculum areas: physical development, health, body care; individual social and emotional development; ethical behavior, standards, values; social relations; the social world; the physical world; esthetic development; communication; quantitative relationships.

Koerner, James D., editor, The Case for Basic Education: A Program of Aims for Public Schools. Boston: Little Brown & Co., 1959, 256 pages

The Council for Basic Education states as its belief that the purpose of education is "the harmonious development of the mind, the will, and the conscience of each individual so that he may use to the full his intrinsic powers and shoulder the responsibilities of citizenship." In support of this statement, the Council directs this book to the layman. In an introductory chapter, Clifton Fadiman defines education as a combination of "socialization and systemization" and the job of the schools as the "transmission and reappraisal of tradition." He argues that basic education includes certain "master subjects" which once learned make it possible to learn other subjects. The "master subjects" are those dealing with language, forms, figures and numbers, laws of nature, the past, and shape and behavior of the earth. Although the core subjects will remain the same, what is taught in each of these subject areas will change as the environment changes. In the articles that follow, educators in specialized fields discuss the relevance of their particular area of study to basic education and give specific subject-matter goals which students should attain by the time they have completed high school. The general areas, each covered by two or more specific articles, are: citizenship, history and geography, English composition and literature, languages, mathematics, science and electives. Herbert M. Schwab offers his view in a summary article that the local board of education is best placed to effect wide changes in school curriculum and urges that changes be made in the direction of basic education.

Keppel, Francis, The Necessary Revolution in American Education. New York: Harper & Row, Publishers, 1965. 197 pages

Mr. Keppel, then Assistant Secretary of Health, Education and Welfare, has directed his book to the layman. He argues that the revolution of "quantity" (of providing some education for all), has been met, that the revolution of "equality" of opportunity is in progress, but that a revolution in "quality" of education is essential. The interrelationships among local, state, and federal governments in regard to education are probed. Central elements of education, teachers, curriculum, research development, and instruments of learning are analyzed.

Lindvall, C. M., Defining Educational Objectives. Pittsburgh: University of Pittsburgh Press, 1964. 83 pages

The central problem explored in this book is the definition of educational objectives. Various approaches to the problem are offered in papers originally presented to a conference sponsored by the Regional Commission on Educational Coordination and the Learning Research and Development Center at the University of Pittsburgh in 1963. In an introduction, C. M. Lindvall emphasizes the fundamental importance of clearly stated objectives for any educational undertaking. He firmly believes the global statements of educational purposes often given by educators and government officials have little relevance to curriculum planning or to classroom instruction. Only specific statements of desired behavior of students at the end of a given program can be meaningful. The contributors, including Robert M. Gagne, Robert Glazer, David R. Krathwohl, suggest various methods of establishing goals for specific situations such as for programmed learning and curriculum building. In each case, the authors state that the goals must be precisely defined and must be stated in a way that permits the measurement of the attainment of the goal. For a goal to be useful, there must be a way to determine whether or not it has been met. Ralph W. Tyler summarizes the importance of clear and specific definitions of educational objectives.

Minnesota University Graduate School, Social Science Research Center, Perspectives on the Conant Report. Robert H. Beck, editor. Minneapolis, Minnesota: University of Minnesota, 1960. 103 pages

James B. Conant's book, The American High School Today, provided the basis for the 13th Annual Public Discussion on Problems of Current Interest in the Social Sciences and of Particular Interest to the Citizens of Minnesota. Twenty professional educators probed various aspects of the Conant Report, its recommendations, its value to schools throughout the country, and its relevance to the schools of Minnesota in particular. The 21 specific recommendations of the Conant Report are summarized in an appendix.

National Education Association Commission on the Reorganization of Secondary Education, "The Cardinal Principles of Secondary Education", Board of Education Bulletin, 1918, No. 35. 32 pages

The focus of this report is on the goals of education in a democracy.

The Commission, appointed by the N. E. A., sets forth its beliefs that education should develop the special abilities of each individual and that these abilities should be used to promote the well-being of fellow citizens and the society. The specific objectives leading to these larger goals are listed as follows: 1) health, 2) command of fundamental processes, 3) worthy home-membership, 4) vocation, 5) citizenship, 6) worthy use of leisure, 7) ethical character. Each of these aims is discussed briefly. A section of the article is devoted to the "specializing functions" (centered on specialized training), and "unifying functions" (emphasizing common ideas and ideals of the society) of secondary education. The Commission argues for the standard secondary school of the comprehensive type and for full or part-time universal education to the age of 18.

National Education Association Project on Instruction, Schools for the Sixties. New York: McGraw Hill Book Company, 1963. 146 pages

The National Committee of the National Education Association Project on Instruction has directed this summary report of its work toward school officials, faculties and other persons responsible for decisions concerning education. Three additional volumes, Deciding What to Teach, Education in a Changing Society, and Planning and Organizing for Teaching describe the work in greater detail. The Committee was to define the major issues in critical areas of American education and to make recommendations for programs designed to resolve the issues. The issues under discussion were organized into twelve "decision areas." The first two areas, decision-making and research, and experimentation and innovation, are discussed under the heading, "The Right Questions." The second section, "Deciding What to Teach," encompasses seven areas involving what is to be taught to whom, when, in what order and for what purpose. The final three areas, involving curriculum organization, school and classroom organization, and materials, technology and space, were handled under the topic Planning and Organization for Teaching. A final chapter discusses future needs of the educational system and how these needs may best be met. The specific recommendations are listed in summary form in an appendix. References to other reports and articles of the Committee are also appended.

President's Commission on National Goals, Goals for Americans.
New York: Prwntice-Hall, 1960. 372 pages

The purpose of this book was to stimulate discussion by the American public leading to increased awareness and consensus of the goals of the United States. Following the request of President Eisenhower, the President's Commission on National Goals, administered by the American Assembly of Columbia University and supported by several private foundations, requested papers from 14 acknowledged experts in various fields. Drawing upon these essays, the Commission formulated goals for domestic and foreign programs. Among the eleven domestic goals outlined are those concerning equality, arts and sciences agriculture, and education. The goals of foreign policy presented are 1) building an open and peaceful world, 2) the defense of the free world, 3) disarmament and 4) the support of the United Nations. The first section of the book presents the report of the Commission to the President. The remainder, actually the greater part of the book, presents the invited essays on various topics under the general headings of American Fundamentals, Goals at Home, and The World We Seek. The goals for education are based on John W. Gardner's contribution, National Goals in Education. He states that the basic goal of our country served by education is equality of opportunity for the fulfillment of the individual. He itemizes 25 recommendations for the educational system of the country designed to further that goal. The areas covered by these recommendations are: the teacher, the student, the curriculum, innovation in education, higher education, education outside the formal system, and the sponsorship of education.

Rickover, H. G., Education for All Children: What We Can Learn From England. Hearings before the Committee on Appropriations, House of Representatives, 87th Congress, Washington, D. C. : U. S. Government Printing Office, 1962, 333 pages

In Admiral Rickover's testimony on education in England before the Committee on Appropriations of the 87th Congress, he holds that the great achievement of the decentrally managed state education system of England is in maintaining uniform national scholastic standards. He examines several attributes of the English educational system which he believes could profitably be adopted in the United States. He argues that all American children must be helped to reach higher intellectual levels. To raise the standards of teaching, he proposes a standards committee, a voluntary accrediting agency, which would set up high criteria and accredit those applying who met the criteria. Samples of classroom work and test questions from English schools are appended.

Rockefeller Brothers Fund, The Pursuit of Excellence: Education and the Future of America. Special Studies Project V of America at Mid-Century Series. Garden City, New York: Doubleday & Company, 1958. 49 pages

The report on education of the special studies project of the Rockefeller Brothers Fund is one of a series concerned with major problems and opportunities confronting the United States in the 1960's. John W. Gardner served as chairman of the Panel on Education and was the principal author of the report. The central theme of the report is that each individual must have the opportunity to obtain the highest quality education so that each may reach his greatest potential. The ideals of excellence must be transmitted and perpetuated at all levels. Needs of the individual, needs of the society, and ways these needs may be met are cited. A section on the educational system contains recommendations for the teaching profession, for curriculum, for the identification of talent for for financing. The utilization and waste of talent are explored and recommendations for increased utilization of talent are presented. The necessity of high motivation and high moral values for our society and their relation to the school are discussed.

United States Department of Health, Education and Welfare, Office of Education, Contemporary Issues in American Education. United States Office of Education Bulletin, No. 2, 1966. Washington, D. C.: U. S. Government Printing Office, 1965. 158 pages

The papers in this collection were prepared by consultants to the 1965 White House Conference on Education and were used as a basis for discussions during the Conference. The President directed the Conference to consider the following question: "How can a growing nation in an increasingly complex world provide education of the highest quality for all of its people?"

To help answer this question, the papers covered a wide range of topics including: vocational education, the role of the states in education, world responsibility, the special student, education of the young, desegregation of schools, higher education, innovations in education, and urban education.

United States President's Commission on Higher Education, Higher Education for American Democracy, Volume I, Establishing the Goals. Washington, D. C.: U. S. Government Printing Office, 1947. 103 pages

In the aftermath of World War II, as thousands of veterans returned to college, President Truman appointed the Commission on Higher Education to examine the functions of higher education and the means by which these functions could best be performed. The final report of the Commission was issued in six volumes, the first of which concerns aims of education. Relationships between education and democracy, international understanding and cooperation are postulated. In a chapter on education for all, barriers to equal opportunity and ways to insure equal opportunity are indicated. Comparisons and contrasts in aims and methods are drawn between general education and vocational education. Specific functions and contributions of several types of institutions of higher education are pointed out. Those studies are the community college, the liberal arts college, the professional school, the graduate school, the research program, and adult education. In concluding, the Commission calls for an education that will develop in students a "democratic dynamic" and at the same time prepare students to contribute toward world order and peace.

III EVALUATION

Ahmann, J. Stanley and Glock, Marvin D., Evaluating Pupil Growth. Boston: Allyn and Bacon, Inc., 640 pages

This text for undergraduate and graduate courses combines an overview of the functions of educational goals with a presentation of the uses and techniques of student evaluation. The authors believe that evaluation should begin with the formulation of goals followed by statements of goals in operational form and then evaluation of students in reference to these stated goals. In the first of four major sections, the role of evaluation and the use of educational objectives as a basis for evaluation are presented. Part II, discusses "teacher-built" tests, while Part III covers test construction and evaluation. In Part IV, existing standardized tests and their uses are reviewed. Interspersed with the text are problems for the student. Appendixes present basic statistical techniques relevant to testing, standardized test publishers and distributors and selected lists of standardized achievement tests, aptitude tests and personal-social adjustment instruments.

Aikin, Wilford M., The Story of the Eight-Year Study. New York: Harper & Brothers, 1942. 157 pages

The Eight-Year Study was a result of efforts of the Progressive Education Association and was financed primarily by the Carnegie Corporation and the General Education Board. The purposes of the Study are to "establish a relationship between school and college that would permit and encourage reconstruction in the secondary school," and "to find, through exploration and experimentation, how the high school in the United States can serve youth more effectively." The planning commission believed that education should utilize all that was known about the way "human beings learn and grow." The statement was made that "the whole boy goes to school; therefore school should stimulate his whole being. It should provide opportunities for the full exercise of his physical, intellectual, emotional, and spiritual powers as he strives to achieve recognition and a place of usefulness and honor in adult society." So that students graduating from participating schools with experimental programs would not be penalized, the cooperation of colleges was sought. Nearly all accredited colleges and universities agreed to waive traditional admission requirements for those students applying from the schools in the study. Thirty schools, all of which wished to make fundamental changes in their educational programs, were selected to represent various regions of the country, various sizes of schools, rural and urban areas, and public and private schools. Each school carried out its own research project. The students were then followed-up in college. It was found that the students of the 30 schools were not handicapped in their college work and were able to assume the responsibilities of a college program. Those coming from schools with the widest curriculum changes reached higher college standing than those against whom they were compared. Conclusions drawn from the Study and implications for secondary education are presented in a summary chapter. An appendix presents the proposal for coordination of school and college work.

American Council on Education, Executive Committee of the Cooperative Study in General Education, Cooperation in General Education. Washington, D. C. : American Council on Education, 1947. 240 pages

Under the direction of Ralph W. Tyler and Ralph W. Ogan, 18 colleges and universities contributed to this study of general education. The purpose of the study was three-fold: to cause desirable change in educational practice, to develop leaders in general education on the various faculties participating in the study, and to promote communication among institutions concerning their contributions to the study and dissemination of the results of the study to all American institutions of higher education. Each college decided what aspects of its general education program to improve, and the methods to be used, and each college conducted its own study. The study was in progress for nealy five years. Workshops, small intercollege committee meetings and regional conferences were held under the auspices of the study. At the conclusion of the study, this Executive Committee Report and its three companion volumes, General Education in the Humanities, General Education in the Social Studies, and Student Personnel Services in General Education were issued. The Report itself comprises topics covering the background of general education, contemporary issues in general education, organization of the study, reviews of the major projects included in the study (the major portion of the book), the hypothetical process of development of a general education program within a college, and a section on the characteristics of general education and conclusions drawn from results of the study. A short section on evaluation is included in this portion. An appendix lists all the projects completed under the auspices of the study.

American Council on Education, Commission on Accreditation of Service Experiences, A Guide to the Evaluation of Educational Experiences in the Armed Services. Washington, D. C. :American Council on Evaluation, Revised edition, 1954. 426 pages

The original guide to evaluation, published in 1944, was intended to help educational institutions in determining the amount and type of credit to be granted for training completed in the armed services. Since the type and range of training offered by the services had changed drastically over the years, a revised edition was prepared by the Commission on Accreditation of Service Experiences of the American Council on Education with the financial support of the Department of Defense. All training programs conducted by each of the services since 1946, whether current at the time of the revision or not, are included in the guide. Each course is described by location, length, objectives and credit recommendations. These recommendations, to be used as guides only, were made by staff and consultants and are said to represent a "conservative estimate of the academic value of the various courses."

Anderson, C. C., and Hunka, S. M., "Teacher Evaluation: Some Problems and a Proposal", Harvard Educational Review, 33, 1963. pages 74-96

Proceeding from findings in the fields of learning theory, personality theory and statistics, the authors suggest measures which would reduce variation in evaluations of teachers due to personality factors of teachers resulting from personality factors of the evaluators themselves. An extensive bibliography is provided.

Archambault, Reginald, "Criteria for Success in Moral Instruction", Harvard Educational Review, 33, 1963. pages 472-483

The author draws heavily from Israel Sheffler's The Language of Education in attempting to define the educational institution's responsibility for moral training. Sheffler distinguishes between "active interpretations" and "non-active interpretations" of morality or norm statements. Acting upon the former, a student would behave according to the morality statement. Acting upon the latter, he would know the moral way but would not necessarily behave according to it. Archambault, modifying Sheffler's postulates, suggests those institutions believing that part of their responsibility lies in moral training should teach students to apply active intellectual processes to questions of morality. This application, he believes, will be conducive to moral conduct but will not necessarily ensure it.

Barton, Allen H., Organizational Measurement and Its Bearing on the Study of College Environments. New York: College Entrance Examination Board, 1961. 82 pages

The College Entrance Examination Board, concerned with non-intellectual elements of students' college success, has published this volume dealing with "contingent factors" characterizing the college. These factors, such as social climate, peer-group relations, etc., are thought to "qualify" the predictive value of high school grades and other predictive measures. From the literature on social organizations, the author has derived a set of categories useful in describing and comparing social systems such as business organizations as well as schools and colleges. The six-fold scheme, not intended as a basis of theory but as a guide to areas of research, permits classification of variables by "external characteristics" including inputs such as economic variables, outputs such as decisions, and environmental variables such as public relations or by "internal characteristics" which include social structure variables, attitudes and activities. Research findings are discussed under the various category headings. The studies reviewed are listed in the appendixes by category for cross-reference purposes.

Cartter, Allan M., An Assessment of Quality in Graduate Education, Washington, D. C.: American Council on Education, 1966. 131 pages

The Commission on Plans and Objectives for Higher Education, under the direction of the author, conducted a survey study of quality of graduate education in 1964. The opinions of 5,367 department chairmen, noted scholars and other faculty members representing 29 fields and 106 institutions were sought. The respondents were asked to select from lists of several terms the one most nearly describing the department in his own field for each school according to these three separate indices: quality of graduate faculty, effectiveness of doctoral program and expected change in relative positions of departments in the next five to ten years. From the information gathered in the more than 4, 000 replies, the author has compiled rankings within each field. The position, training and professional participation of the respondents were also scrutinized. A panel of experts in four fields was chosen to study the selected fields in greater detail and to provide data for reliability studies.

Chronbach, Lee J., "Course Improvement through Evaluation", Teachers College Record, May 1963. Pages 672-683

Three types of decisions for which evaluation is used are said to be course improvement, decisions about individuals, and administrative regulation, each of which requires its own measurement techniques. The author contends it is less important to obtain individuals' precise scores for course evaluation than for decisions among individuals. Furthermore, scores on individual items may be more important than an overall score. He believes the greatest service evaluation can perform is to identify aspects of a course where revision is desirable. Opinion can be used, he states, as a criterion in formative stages of curriculum development to eliminate the most obvious problems but the more systematic analysis should be employed. He discusses the limitations of course content items in the evaluation of courses and the different types of measurement required for different types of transfer of learning situations. He concludes that "... course evaluation calls for description of outcomes. This description should be made on the broadest possible scale, even at the sacrifice of superficial fairness and precision."

Dressel, Paul L. and Associates, Evaluation in Higher Education. Boston: Houghton Mifflin Company, Boston, 1961. 480 pages

Individual chapters of this general presentation of evaluation in the college setting were authored by several contributors. However, Paul Dressel as general author provided the continuity of the volume and co-authored several chapters. The first chapters deal with the nature and purposes of evaluation. The authors contend that evaluation should be built into any program of education. One chapter is devoted to the mechanisms of learning while four chapters relate evaluation to specific subject areas.

The final chapters examine specific evaluation situations such as testing and grading policies, selection and placement of students, evaluation of instruction, and institutional self-examination. Appendixes comprise a short discussion of technical matters relating to evaluation and a tentative outline for an evaluation survey of higher education in Michigan.

Dressel, Paul L. and Mayhew, Lewis B., General Education: Explorations in Evaluation. The Final Report of the Cooperative Study of Evaluation in General Education of the American Council on Education. Washington, D. D.: American Council on Education. 302 pages

The final report of the Cooperative Study of Evaluation in General Education of the American Council on Education summarizes studies of the effects of programs in general education in 19 colleges. The overall aim of the study was to define the objectives of general education, to develop evaluation techniques and to collect evidence on student achievement. Relationships were sought between behavioral goals (rather than philosophic goals) and instructional methods and instruments (rather than course content or subject matter). The first two sections of the book outline the background of the study. Reports covering studies in specific areas of general education (social science, communications, science and humanities) are followed by reports of studies on traits (critical thinking and attitudes). Implications and perspectives conclude the report. Lists of materials developed for the study and of objective tests in communications are appended.

Eckert, Ruth E., Outcomes of General Education. Minneapolis, Minnesota: University of Minnesota Press, 1943. 210 pages

When the experimental General College of the University of Minnesota was established in 1932, a program of evaluation was envisioned as part of the general development of the College. The College was designed to provide students of average ability with two years of liberal education to prepare them for problems they would encounter after college. In preparation for six years, the evaluation studies began in 1938. It was felt that the success of the college curriculum ought to be judged by "products," changes in thinking and behavior of all the students enrolled for even a short time in the program, and not by "process," the design of the curriculum itself or the skill of individual teachers. During the preparatory years, surveys determined the abilities , interests, attitudes and family backgrounds of students and major problems, activities and needs of former University of Minnesota students. The first step in the evaluation itself was to define the goals of the General College program. Studies were then carried out to provide clear descriptions of the General College students, the programs they chose and the degree of success in those programs. The major concern of the study was to identify the changes in the information, attitudes and adjustment in the course of the General College program. Student and faculty attitudes toward the program itself were also measured. Follow-up studies involved drop-outs and those who transferred to other divisions of the University. The report discusses each phase of the evaluation with chapters on such topics as progress in the General College, readiness for continued learning, and vocational readinesss. Strengths, weaknesses and recommendations for steps to counter weaknesses are presented for each topic.

Flanagan, John C., Design for a Study of American Youth. Boston: Houghton Mifflin Company, 1962. 240 pages

Project TALENT was designed to facilitate the identification and development of talent and to find ways to increase the probability of educational or vocational success following high school. The investigators hoped to discover the characteristics of high school students that would demonstrate the determinants of success during school years and later. This book, the first in a projected series of four, details the background and methods of the study which involved students from nearly 1,000 schools in all parts of the United States. The selection of the sample, the development of the tests, and the collection of data are all recounted. The tests covered areas of educational achievement, vocational choice, personality factors and personal history. Procedures for gathering related data from school administrators are also presented. Plans for analysis of the data and follow-up studies are detailed.

Furst, Edwaurd J., Constructing Evaluation Instruments. New York: Longmans, Green and Company, 1958. 334 pages

The author, concerned with the evaluation of student progress, presents the basic problems of evaluation to the classroom instructor and provides an introduction to methods of test construction. He defines the major components of a "well-planned program of education" as educational objectives, learning experiences and evaluation procedures. The concepts are interrelated as the objectives serve as the basis for developing learning experiences and evaluation procedures which then help to clarify the objectives. The basic problems explored in Part I include stating the desired outcomes, determining the specific behavior corresponding to general objectives, and obtaining information of progress relevant to the desired outcomes. Part II reviews methods of test item construction and administration, scoring and revision of tests.

Great Britain Secondary School Examinations Council Committee on Curriculum, Curriculum and Examination in Secondary Schools. London: His Majesty's Stationery Office, 1943. 151 pages

The general purposes of education in Great Britain and the principles underlying the educational curriculum are presented. Special problems of creating a curriculum for lower schools and for higher forms are dealt with. A chapter is devoted to recommendations for each of 13 areas of the curriculum.

Guba, E. G. and Getzels, J. W., "Personality and Teacher Effectiveness: A Problem in Theoretical Research" Journal of Educational Psychology, 46, 1955. pages 330-344

The main tenet of this article is that research based on a theoretical framework is superior to fact-finding or normative approaches in studying the relationship between personality factors and teacher effectiveness. One such research effort is outlined and its advantages listed.

Haggerty, Melvin E., The Evaluation of Higher Institutions, Vol. II The Faculty. Chicago: The University of Chicago Press, 1937. 218 pages

The Committee on Revisions of Standards was established in 1929 by the Commission on Institutions of Higher Education of the North Central Association of Colleges and Secondary Schools. The Committee was asked to evaluate the standards and accrediting procedures used by the North Central Association and to develop new criteria for the measurement of institutions. Over 50 institutions contributed comprehensive information compiled from questionnaires and tests. Volume II primarily concerns the measurement of faculty excellence.

The first task was to create a criterion of general educational excellence. A preliminary attempt was made to build a criterion based solely on the quality of graduates as determined by alumni achievement. Such a criterion would disregard such factors as faculty, curriculum, endowment, etc. For lack of comparable data from the different schools, the attempt was unsuccessful. The final criterion evolved from (a) several separate ratings of general educational quality, (b) composite rankings based on tests measuring reading comprehension, natural science knowledge, mathematics knowledge, knowledge of French and German, etc., and (c) ranking of institutions based on work of their graduates in graduate school. The next step was to "evaluate certain factors in and about the total faculty that appear to parallel other marks of institutional quality." An index based on (a) faculty competence as judged by such factors as highest degree obtained, publications, memberships in professional associations, attendance at professional meetings, etc., (b) faculty organization, including faculty and student ratio, administrative obligations, etc., and (c) conditions of faculty service including starting salary, retirement, benefits, etc., was set up. The relationship between the faculty index and the criterion was determined and found to be inadequate to warrant the use of the index as a sole measure for the accreditation of colleges. Other volumes of the series deal with various other aspects studied, e.g. finance, plant, etc. Of all the variables studied, faculty competence had the lowest correlation with the criterion, and faculty organization the highest. It was found that none of the variables studied, or even the combination of these variables, was sufficiently highly correlated with the criterion to warrant use as a determinant of accreditation. Therefore, the desired goal of specific standards was abandoned and general statements covering all aspects of schools related to academic excellence were established. Those statements regarding the faculty are recounted here.

Hyman, Herbert H., Wright, Charles R., and Hopkins, Terence K., Applications of Methods of Evaluation, Four Studies of the Encampment for Citizenship. Berkeley and Los Angeles: University of California Press, 1962. 392 pages

The authors seek to clarify and work toward the solution of problems in methods of evaluation through a case study of research done over a four-year period to evaluate a social program, The Encampment for Citizenship. Descriptions of the campers and the camp program provide the background for the discussions of the changes in attitudes, values, etc. brought about by the program. Effects of the program, lasting through the return home and into college, are studied. Long-range effects of the 1955 encampment and earlier encampments are also discussed. Appendixes include attitude scales, questionnaires, tables, and descriptions of statistical methods.

Lindquist, E. F., editor, Educational Measurement. Washington, D. C.: American Council on Education, 1951. 819 pages

The Committee on Measurement and Guidance of the American Council on Education, with support from the Grant Foundation, provided the impetus for this volume edited by E. F. Lindquist. The work was designed as a handbook and textbook covering both theory and techniques of educational measurement. Recognized experts in each of several areas of psychological testing contributed chapters covering their specific areas of competence so that a wide range of developments in testing could be examined thoroughly. The chapters are organized into three major divisions: Functions of Measurement in Education; Construction of Achievement Tests; and Measurement Theory.

National Advisory Mental Health Council, Evaluation in Mental Health. Washington, D. C.:Public Health Service, U. S. Department of Health, Education and Welfare, 1955. 292 pages

Nearly 1,000 bibliographical items are annotated in this review of studies concerned with evaluation. The items were selected either for their contributions to general theory and methodology or to mental health activity areas, including: community organization, administration, professional personnel, education and information, preventive effects of programs, factors influencing individual mental health, and diagnostic, prognostic and treatment procedures. The studies of mental health activity areas were selected for their presentation of the methods of evaluation employed rather than for the reported results of the studies. The areas covered by the bibliography are discussed with particular emphasis on theoretical aspects of evaluation. The major portion of the volume consists of the study summaries which are presented by sections corresponding to the areas discussed in the preceding text.

National Education Association, Profiles of Excellence. Washington, D. C.: National Education Association, 1966. 126 pages

Profiles of Excellence was designed as a tool for the evaluation of school system operation and was based on the premise that certain aspects of a school's operation serve to encourage and support educational excellence. The instructional program itself is not considered in this scheme. The guide was devised by the Office of Professional Development and Welfare of the N. E. A. and was pre-tested in 11 school systems in various parts of the country. The criteria employed are based on results of related research efforts and professional judgments and practices. Recommended techniques for the use of the criteria preceed the section of specific evaluation materials. Areas of educational system for which materials are provided are: the educational program; administration; Board of Education; staff personnel policies and procedures; professional compensation; the school plant; district organization, finance and business administration; and local professional association. Forms for summary profiles are appended.

National Study of Secondary School Evaluation, Evaluative Criteria. Washington, D. C.: National Study of Secondary School Evaluation, 1960. 373 pages

The major concerns of the National Study of Secondary School Evaluation, formerly the Cooperative Study of Secondary School Standards, have been to determine the traits of good secondary schools and to develop improved methods of evaluation of schools. Members of The Study are representatives of six regional accrediting associations and advisory members. To take into account equal quality among schools with quite different purposes, The Study acts on the principle that "a school should be evaluated in terms of what it is striving to accomplish (its philosophy and objectives) and in terms of the extent to which it is meeting the needs of the students who are enrolled and for whom it is responsible." The Study has published three editions of Evaluative Criteria, a guide for the evaluation of schools. Studies to assess the value of the evaluation program have been carried out over the years and appear to confirm the effectiveness of the program. The handbook contains a manual describing the background of The Study, the purposes of evaluation, and the procedures to be followed in an evaluation, as well as work sheets for actual evaluations. The manual instructs the schools to determine the basic philosophy of the school and then to translate this philosophy into specific objectives. It is according to both the philosophy and stated objectives that the evaluation is then to proceed, using the work sheets provided for each of 19 subject areas and 7 areas of administration, service, and nonacademic elements of the program. Within each subject area, check lists and evaluations are to be completed for course organization, nature of offerings, physical facilities, direction of learning (instructional staff, instructional activities, instructional materials and methods of evaluation), and outcomes. Appendixes provide forms for statistical summaries and graphic summaries of evaluation.

National Study of Secondary School Evaluation, Evaluative Criteria for Junior High Schools. Washington, D. C.: National Study of Secondary School Evaluation, 1963. 350 pages

The junior high edition of Evaluative Criteria was prepared to meet a demand for materials specifically geared to junior high programs and is a modification of the 1960 edition of Evaluative Criterial for secondary schools. The format is identical to that of the parent edition but the subject areas covered by the work sheets more accurately reflect the purposes of the junior high school. Thirteen subject areas and seven administrative, service, and nonacademic aspects of the program are covered.

Remmers, H. H. and Gage, N. L., Educational Measurement and Evaluation. New York: Harper & Brothers, 1943. 580 pages

The authors devote the first section of the book to the question, "What should be evaluated?" Discussions of psychological theory relevant to the testing of various aspects of students' behavior and abilities are included in this section. The remainder of the book is concerned with test construction, administration, interpretation and application of test results. Two chapters on statistical theory useful for the understanding of psychological tests are also included in this section.

Seiler, John A., Systems Analysis in Organizational Behavior. Homewood, Illinois: Richard D. Irwin, Inc. and the Dorsey Press, 1967. 219 pages

The author takes as his starting point the idea that organizational behavior occurs as a result of a system of interdependent forces. He presents an analytic framework for the analysis of organizational systems based on human behavior within the organization. He states that his model, though primarily concerned here with business organizations, is applicable to any type of formally organized institution. In the framework, he distinguishes between environmental forces and the organizational system. He postulates that the environment places restraints upon the organizational system and that the organizational system makes choices based upon forces in the environment. "Inputs" within the system -- human, technological, social and organizational - have a functional relationship with "Actual Behavior" - activities, interactions and sentiments - which in turn have a functional relationship with "Outputs" - productivity, satisfaction and development. Feedback facilitates relationships among the factors. The first two chapters of the book explore elements of systems and describe the analytic framework. The latter chapters apply various parts of the framework to case studies.

Thorndike, Robert L., and Hagen, Elizabeth, Measurement and Evaluation in Psychology and Education. New York: John Wiley & Sons, Inc., 1955. Revised edition 1961. 602 pages

Throughout this text on the fundamentals of testing, the authors emphasize the necessity for precise definitions of the objectives being measured. The text is intended for teachers and other education personnel as well as for college students majoring in psychology. Classroom tests, achievement tests, ability tests, interest tests and personality tests are all discussed along with basic statistical methods necessary for interpretation of test results. Uses of tests in educational and vocational guidance and in industry are indicated. Selected standardized tests and their sources are listed in the appendix.

Tucker, Ledyard R., Formal Models for a Central Prediction System, Psychometric Monographs, No. 10, 1963. 61 pages

To explore the problem of predicting college academic success from high school grades and supplementary data, Educational Testing Service established a committee to examine the technical aspects of a central predictive system. As part of that effort, and with support from the College Entrance Examination Board, the author investigated possible mathematical structures of such a system. The factors he wished to include in such models were high school student test results and grades from each of several schools and subsequent college grades of these students at a number of colleges. He notes that previous studies traditionally have been based either upon students from one high school at several colleges, or students from several high schools at one college. Three models are presented and their uses explored and compared.

Tyler, Ralph A., Gagné, Robert M., and Scriven, Michael, Perspectives of Curriculum Evaluation. Chicago: Rand McNally & Company, 1967. 102 pages

In the opening chapter, Robert Stake notes the inadequacies of present evaluation techniques and enumerates major needs in the evaluation field. He outlines the history of the Committee on Curriculum Evaluation leading to the publication of this monograph. A major point made by Ralph Tyler is that little or no notice is taken of the effect of changes inaugurated in traditional evaluation studies on the pre-existing educational structure. In addition, he contends, innovations in education make traditional use of the success in college as criteria for judging the worth of high school educational programs outdated. Robert Gagné discusses curriculum in terms of a "hierarchy of content units" wherein each unit depends on the mastery of the previous unit. The curriculum could be, then, a sequence of content units, the learning of each being a single act if the previous unit has been mastered. In his paper, Michael Scriven distinguishes between formative evaluation, carried out at one or more intermediate stages of curriculum development, and summative evaluation, carried out at the end of a project. He also makes a distinction between intrinsic evaluation, concerned with specific features of the teaching instrument or new method being adopted, and pay-off evaluation, concerned with effects of the method on the student. Stanley Ahmen summarizes the previous papers and suggests that try-out types of studies should be emphasized rather than final decision ones. The former being more tentative can employ less sophisticated measurement techniques. He also indicates his belief that local school officials should pay less attention to universal summative evaluation studies and more to local summative ones.

UNESCO, (Klineberg, Otto, et al) Evaluation Techniques, International Social Science Bulletin, Vol. VII, No. 3, 1955

This issue is devoted mainly to a collection of discussion papers on the methods and technology of evaluation gathered for the International Social Science Council and published here in shortened versions. In his introduction to the papers, Otto Klineberg states that "evaluation is a process which enables the administrator to describe the effects of his programme, and thereby to make progressive adjustments in order to reach his goals more effectively. It is important to note that it is not restricted to application at the end of a programme, but involves periodical investigation at many stages." He discusses possible uses of evaluation and the contributions of social science to evaluation procedures. Section I, Methods and Results, includes a paper on the nature of evaluation followed by papers on evaluation in each of several fields such as intergroup relations, exchange of persons, and mass media campaigns. Section II is a report of the meeting on criteria and techniques of evaluation of technical assistance for economic development sponsored by UNESCO in 1954. Topics covered include the purpose of evaluation, materials and methods of evaluation and the evaluational process related to technical assistance.

D. APPENDICES

2. Evaluation of Educational Goals

Joseph Lennards

EVALUATION OF EDUCATIONAL GOALS

Joseph Lennards

September, 1966

In all Western industrialized societies the educational system is under critical review at the moment. There is disagreement and confusion about

1. what the goals are which the school should try to accomplish;
2. what contribution we can legitimately expect from the school.

In many European countries the uncertainty is increased by the traditional opposition of educational theorists to empirical research. The philosopher is still king and the goals and methods of education seem to be evaluated more on the basis of the desirability of their ideals than on the basis of their empirical validity and consistency. An empirical assessment of these theories could contribute considerably to the "de-ideologization" of much of the present educational thinking. Such an assessment of educational theory would involve the following steps:

A. Identification of the empirical content of the theory.

1. a) The (most of the time) wholistic formulated goals of the theory are to be identified and are to be broken down into specific sub-goals.
   b) These sub-goals should then be ordered according to their mutual relationship. Is the one precondition for the other? Is there a processual development implied?
   c) What is presented as the psychological and sociological foundation of these goals and of their mutual relationship?
2. a) What are the methods and educational means recommended for achieving any specific goal?
   b) What specific contribution does one expect from any specific means?
   c) What is presented as the psychological and sociological foundation for these expectations?
3. a) How does one know whether the goal is achieved? What behavioral or attitudinal characteristics are supposed to indicate a successful education? In what population are these to be found?

b) What is the psychological and sociological foundation for proposing these operational criteria?

B. Empirical evaluation of the theory.

1. On the basis of our psychological and sociological knowledge what can we say about
   a) the internal consistency of the goals (can we be "excellent" and "equal" too?)?
   b) the external consistency of the goals in the daily life of children, adolescents and adults. Are these goals acceptable in the society?
2. On the basis of our psychological and sociological knowledge what can we say about the postulated connection between goals and means?
   a) Are the proposed means capable of producing the expected result?
   b) Are they internally consistent (mutually supporting or interfering with one another)?
   c) What is their external consistency? Can they be institutionalized? More specifically, what do these means imply about the role of the teacher (authority, competence), the availability of such teachers, the role of the child (intelligence, motivation), the availability of such children and the allocation of economic resources?
3. On the basis of our psychological and sociological knowledge what can we say about the proposed operational criteria?
   a) Of what particular goal is this an operationalization? Is it a valid operationalization? Do the operational criteria cover all the goals?
   b) What are the appropriate measurement instruments to test the effectiveness of the educational theory?
   c) What do the results of our measurement show?

Although it is important to keep the above model in mind as an ideal, it will often not be possible to apply it. Many educational theories are defined in a vague way (the "cultivation of a liberal mind") with no or little specification of the means required, of the exact relationship between goals and means and of the operational criteria. Obviously, the acceptance and popularity of these theories depends less on their empirical soundness than on their ideological appeal to certain subgroups in the population. Demonstrating this lack of empirical foundation in educational thinking is certainly an important task for the sociology of education, but for our purposes it may be better to view it as a secondary one.

Our interest lies in finding out whether educational systems reach their goals. The word "their" in this context refers - if I understand it correctly - to society as the frame of reference. The goal statements to be found in the educational literature have, however, mostly a different system reference: not society, but the school as an autonomous institution. This applies especially to the European countries, where the professionalization of teachers and educators is much more advanced and the formulation and shaping of educational policy is greatly influenced by their ideals and interests. Much of the current debate in these countries centers exactly around the question whether the educators are not holding up an ideal that is no longer adjusted to the requirements of a modern industrial society. By taking the proclaimed aims as focus of our investigation we may discover that schools achieve THEIR purposes, but this does not give us sufficient assurance that the educational system also reaches its societal goals. Secondly, even if we could find in the educational literature goal statements with a clear societal reference, this only gives us the manifestly perceived goals. From the point of view of society, the educational system may have additional goals which are only dimly perceived. A third reason for not concentrating, in first instance, on the officially stated educational goals is that such a procedure makes it very difficult to develop a model which is applicable cross-culturally. If we want to compare the relative effectiveness of educational systems in reaching their goals, then these goals must at least be stated in terms of common categories in order to make meaningful comparisons possible.

These categories should express what is common to the role of the school in industrialized, democratic countries and should, at the same time, be specified in such a way that the differences in emphasis with regard to these common objectives can be adequately investigated. Almond's and Verba's "The Civic Culture" provides a useful negative example of what is meant here. In this book the definition of what constitutes democracy is derived from the American experience (participation in voluntary organizations, etc.) This operationalization is then applied to the countries in the sample. It is, of course, not surprising, that the United States appear to have a highly developed "Civic Culture" and that Mexico ranks lowest. As far as the U.S.A. is concerned, this result is tautological (output = input) and, as for the other countries, the comparison is often meaningless. They are confronted with an ideal operationalized in such a way that it does not reflect sufficiently their definition as to what constitutes a democratic climate or an ideal that is not applicable in their particular environment (given e.g. the literacy rate in Mexico). In our search for common categories we must try to avoid imposing on the data a framework that lacks objective relevance or that only partly expresses the phenomenon in question and neglects certain other dimensions. In order for a comparative model to be applicable meaningfully it should contain common categories and, at the same time, it should incorporate a sufficiently detailed specification of the different and alternative ways in which this common dimension can be expressed. To give a concrete example: a common goal of the school in democratic societies is education for citizenship. What this involves depends a.o. on the definition of democracy prevalent in a particular society. If democracy is interpreted as representative democracy (competition between representative elites), then political socialization for the majority will aim at creating the passive citizen whose influence is confined to passing judgement, at regular intervals, upon the activities of the elite minority. If democracy is defined as direct democracy, the role of the citizen will be a more active one and citizenship training should, accordingly, be more intense and comprehensive.

If this opinion about the necessity of developing a model which contains both common goals and the possible interpretation of these goals, is accepted, certain consequences follow as to the long-term strategy of the project. Although the final model will be applied to the American educational system first, it should not be based on an analysis of the American data only, but on a comprehensive survey covering all the countries to be included. In order to be able to compare the "products" of the American school system with the "products" of the European system all the possible variables should be included from the beginning. Thus, although we may have no reason to expect that there will be marked feelings of superiority and inferiority among the pupils of different streams in the American High School, still it is crucial to include this variable, because, otherwise, we have no way of testing which system is less elite-oriented than the other. Because of this requirement it may be advisable to have the meeting with the European experts rather early in order to obtain the fullest possible specification of the relevant dimensions.

What are the common goals of the school in an industrialized, democratic society? The following four structural changes seem to me to be the most important ones in determining the relationship between school and society.

1. Loss of functions for the family.
   The present family as a system of personal relationships can only make a partial contribution to the socialization of the young. As Parsons has indicated ("The School Class as a Social System") it is the function of the school to make possible the transition from a life bounded by the family and its intimate relationships to a life of participation in which other kinds of relationships (especially, impersonal

ones) are found. The goal of the school is to extend the social skills of the child and to orientate him towards the larger society. This orientation includes political socialization and occupational socialization. The following three changes have made the role of the school with regard to these three sectors crucial.

2. The changes in the stratification system which can be summarized under the heading: the extension of civic, political, and social citizenship (T.H.Marshall). The emphasis on the equal political rights (and, consequently, duties) of each citizen has introduced citizenship training as a necessary element in the school curriculum. The attempt to reduce the influence of hereditary privileges and to provide equal opportunities to all citizens has brought about a fundamental change in the role of the school with regard to the stratification system. Instead of exercising merely a differentiating function the school has now become the main arbiter of social stratification and the principal mechanism for the discovery, training and allocating of talent.
3. The changes in the structure of production have made the role of the school as the main provider of manpower also an economic necessity (in addition to being a democratic right). Economic strength and development depends on technology and technology, in turn, depends on an educational system which can develop and support the technology. This involves two things:
   a) The school must be succesful in preventing the loss of talent and in cultivating excellence;
   b) The school must inculcate values, skills, and attitudes which are in accordance with the technological character of the occupational world.
4. The changes in the structure of consumption have opened up opportunities for social and cultural participation not available to the majority of the population before. This equalization of the rights of each citizen to cultural development requires an extension of the role of the school. Its

task includes now also raising the cultural level of the whole nation.

We may summarize the above by saying that in an industrialized, democratic society the attainment of the economic, political, social and cultural goals is greatly determined by the effectiveness of its educational system.

In reading through the European literature on the necessity of school reform I have discovered these same themes. The European system is criticized because of its undemocratic character both in its selection procedures (see 2. above) and in its distinction between elite and mass education (see 4. above). At the root of the European school system lies the distinction between a school for the elite and a school for the people. The former aims at providing a "liberal" education, that is, the training of the intellect to attain knowledge for its own sake. The latter provides a "servile" education, that is, the transmission of knowledge and skills in so far as needed for the production of goods or the performance of services. This discrepancy between democratic ideal and actual educational practice has been recognized and in all Western European countries steps have been undertaken to reduce the discrepancy. But, these measures are aimed primarily at democratizing the selection procedures, not at bridging the gap between elite and mass education. The present educational provisions are still based on a dichotomy between a liberal and a vocational orientation. The school leaving age has been raised and the amount of educational exposure has been increased by extending and elaborating the people's schools rather than by integrating them with the elite's schools. Not only the requirements of democracy work against this pattern, the needs of industry also ask for change. The critics demand adaptation of the goals of secondary education to the needs of economic expansion (3b above). The dichotomy between humanism and technology, which has its root in Greek thought, has today become anachronistic and a "new humanism" must be formulated. As for the vocational schools, they also are guided by out-moded training models, namely, that of the craftsman or the artisan, who learns a particular job and practices this the rest of his life. Modern vocational training should be less vocational and less specialized. It should provide a basic type of general education and lay more stress on general occupational and working qualities in order to be useful in a society where technological developments create continuous occupational shifts.

The content-analysis of the European progressive literature on school-reform shows that the same themes are stressed as in the sociological literature. The goal of the school is to socialize:

1. the future producer;
2. the future citizen;
3. the future culture consumer;
4. the future adult.

I realize that by emphasizing only these goals several objectives are omitted which figure prominently in the American educational literature on the role of the school. But, if we want to develop a model which is applicable cross-culturally, we must limit ourselves to what is the common core of the educational task. I see no objective reason for including e.g. socialization of the future family member, neither is there a subjective reason. In fact, in most European countries (especially, in France) the parents would object to this expansion of the task of the school. It is important to keep in mind the historical background of the American emphasis on life-adjustment, namely, the necessity to assimilate the children of immigrants into a new culture. Several of the proclaimed aims of American education can be explained as a response to this particular situation rather than as a necessary reflection of the role of the school in an industrialized society. Although the goal-model is less comprehensive than an analysis of the American literature would suggest, it postulates, at the same time, more objectives than can be found in the traditional European literature on the subject-matter. Education is not the transmission of culture for its own sake, but preparation for life in the context of an industrial society.

A second characteristic of the above model is its sociological focus. The emphasis lies on role- socialization, that is, the transmission of information concerning the relevant values and norms, the learning of the skills and aptitudes required for successful performance and, thirdly, the inculcation of commitment to the role. All these processes presuppose the presence of certain psychological facilities and predispositions (the ability to think clearly, the ability to emphasize, etc.) and it is also the task of the school to develop these. There is a "psychological readiness" for socialization, which, in the above scheme, is taken for granted in order to keep the problem within manageable bounds. Still, in order to obtain the information required by the model, we have to rely on psychological achievement tests, if we want to do a complete job. It is questionable whether these tests are available or can be developed within reasonable time. (It took years before experts were able to develop a test for mathematical achievement which could be applied cross-culturally).

With these limitations in mind the following list is offered as a tentative and incomplete specification of the content of the above goals.

1. Producer.
   a) The school should ascertain the child's particular abilities and skills, not just his intellectual ones, and provide him with this information.
   b) The school should provide the child with a knowledge of the occupations compatible to his particular talents.
   c) The school should provide the child with knowledge of the content and prospect of the work (working hours, promotion chances, training requirements).
   d) The school should provide the child with the knowledge of the general occupational structure and of the way in which the occupations are available to him fit into this structure.
   e) The school should provide the child with the skills required for a successful performance on the occupational level of his choice.
   f) The school should provide the child with the motivation to exercise his talents. This includes, where applicable, the motivation to raise the thresholds of his aspirations.

g) In order to assure role-commitment the school should provide the child with a realistic appraisal of the possibilities for human development offered by his particular occupation.

Some possible ways to test the attainment of these goals are:

ad d) Give the pupils a list of selected occupations and ask them to indicate the level of education required for each. Then ask them to rank these occupations according to their prestige.

ad f) and g) Morris Rosenberg in "Occupations and Values" has developed a list for criteria of career choice which could well be applied here. First, ask for the pupils' opinions on the most important criteria for them and then, have them indicate what jobs they think to take and whether, in their opinion, their occupation will satisfy most of the criteria given earlier.

I suspect, that in those countries where the "liberal occupations" are still considered to be the most ideal ones, there will be a discrepancy between criteria for career choice and possibilities of pursuing them in a particular job. This is indicative of the gap between humanism and technology. This can also be tested in another way, namely, by comparing the career choices of pupils in the elite school. Are these only humanistically oriented? It would be advisable to take only the choices of the top ten percent, because the less good students may take up technological jobs since there is no other choice for them.

2. Citizen.

In view of the different interpretations of democracy in the countries to be considered it is necessary to develop a typology which covers the range from active-passive citizenship. I may refer to the Litt article(see bibliography) and to Habermas' book (idem).

A distinction should be made between "nation" and "state". Political socialization is oriented toward both and these two objectives may be contradictory at times. Creating active citizens means encouraging critical attitudes and readiness to work for the obtainment of power if values are no longer considered to be acceptable. Political socialization also means the inculcation of a sense of national identity. If patriotism is too strong, however, a critical attitude may be discouraged.

The common core of political socialization is

a) knowledge of the rules of the democratic game.
b) knowledge of the institutional way in which power is organized and can be influenced.
c) knowledge of what constitutes the particular character of the nation and of its position in the larger world.
d) an attitude of tolerance; the willingness to hear everybody's opinion as potentially important; the ability to sort out information critically and the ability to accept compromises.
e) a feeling, that the power structure can be influenced and that it should be done by organizing for democratic action.

3. Culture consumer.

Here socialization involves:

a) Knowledge of the opportunities available for a meaningful spending of leisure time. This is, knowledge of what particular tastes and sensibilities can be satisfied by what kind of recreation.
b) the possession of artistic and literary taste. We would like to include also, as a requisite for the enjoyment of cultural opportunities, the ability and the readiness to continue learning.

The education for cultural enjoyment may come in conflict with the education for a meaningful productive life where too much stress is laid on "Culture" as the highest fulfillment of human life.

4. Adult.

Many of the desired aspects of the adult are listed already under the above headings. The most important general characteristic of the role of the adult in an industrial, democratic society is probably: the ability to tolerate change and ambiguity and the readiness to accept progress based on achievement in the sciences and other fields of knowledge. It is difficult to give a more detailed description of the ideal man in the industrial society. The critics of the European educational system are very good in pointing out why the system is no longer adequate, but they are less articulate in specifying what the new image of man should be. (see, Hoffman's remarks on this question).

After having specified further the variables to be included in the goal model the next task will be to identify what the role of each type of school in the educational system is with respect to these goals. In other words, it is necessary to specify what contribution to the total finished product is expected from each type of school. A second question to be considered is whether the educational system has the complete responsibility for achieving these goals or whether other institutions are supposed to take over part of the socialization process. In several European countries e.g., there is an apprenticeship system. Obviously, industry also takes some responsibility for socializing the producer. This raises the larger problem of the influence of non-official educational institutions, like the home, the community and the mass-media. Unless adequate ways are found to control for these variables, it is impossible to decide whether the successes of failures of the school are really due to the working of educational factors only.

In addition to a description of the goals to be achieved by the educational system of an industrial, democratic society we need a description of the means used in pursuing these objectives. The purposes of this means-model are, in first instance, to provide a descriptive categorization of the relevant aspects of the educational structure. The effects of these structural arrangements on the achievement of the above goals will be indicated, wherever possible. Often the postulated effects are merely research-hunches, but this does not affect their usefulness. The development of an explanatory means-model is the intended end-product of this research project and thus, empirical evidence will show whether our hypotheses were valid or not. The means-model will be presented in a schematic form:

A reading of the review articles in Gage N.L. "Handbook of Research on Teaching" - the most recent and competent compilation of our knowledge in this field - will quickly reveal that our knowledge of the social and psychological factors affecting the learning process is both scanty and mostly inconclusive. The methodological design of many of the studies is weak and the neglect to make certain basic, theoretical distinctions makes it very difficult to compare the results of the studies. So it is, for instance, important to distinguish between cognitive learning (knowledge and skills) and non-cognitive learning (values and motivations) and to realize that what contributes to the achievement of the latter goal does not necessarily contribute to the achievement of the former one also. Cognitive learning is mostly measured in terms of grades or test scores and given this operationalization the experiments only measure the effects on scholastic achievement rather than on intellectual development. The effects of the independent variable are sometimes measured in terms of individual change, sometimes in terms of group change and, of course, this makes the results incomparable. From our readings, we have selected the following variables:

I. Teachers.

a) Social background. Are they mostly upward mobile, stationary or downward mobile?

1. The upward mobile teacher (strong ambition, deferred gratification and hard work) may find it hard to deal sympathetically with pupils which are not ambitious and hard-working. He is probably more successful with middle class and ambitious pupils and impatient with unambitious and working class pupils.
2. The stationary teacher may have chosen teaching for positive reasons: intrinsic enjoyment. He will probably be open to more values than just ambition and achievement.

See: Becker, Howard S. "Social-Class Variations in the Teacher-Pupil Relationship". Journal of Educational Soc. 1952. Vol. 25, pp. 451-465.

Himmelweit, Hilde T. "Socio-economic background and personality," International Social Science Bulletin, Vol. VIII, pp. 29-34 found the upward mobile ~~students~~ teachers most authoritarian.

b) Prestige in the larger community

1. the less the prestige of the teacher in the larger community, the less his moral authority over his pupils (Durkheim "Moral Education"). This will affect, especially, the motivational component of learning.

c) Competence. Is related to teacher's training.

1. The greater the cognitive competence of the teacher, the greater the gain in cognitive learning of the pupils. See: Anderson, K.E. "A Frontal Attack on the Basic Problem in Education: The Achievement of the Objectives of Instruction in Specific Areas". J. of Exp. Educ., 1950. Vol. 18, pp. 163-174.

   The greater the number of sciences courses taken during college education, the greater the rate of student improvement in these courses.

2. The wider the area of his perceived competence, the wider the area of influence of the teacher.

   See: Hovland C. and Kelley H. "Communication and Persuasion". New Haven, 1953 Ch. 2.

   College students responded more favorably to opinions from experts in the field. This "credibility" concept is important, as suggested by Boocock, Sarane S. "Toward a Sociology of Learning". Soc. of Educ. 1966. Vol. 39, p.8. It may explain "...why so few teachers are really effective in shaping students' views on politics, morality, and such - they are simply not perceived as "experts" in these areas". Perceived competence is, of course, related to prestige and this, in turn, is determined by factors like the degree of professionalization of the teaching profession.

II. Content of teaching.

a) Explicitly. To what extent gives the curriculum information about the above goals or is it aimed at cultivating the skills and aptitudes required?

b) Implicitly. An analysis of the textbooks used in history would be useful for discovering what characteristics are presented as being peculiar to this nation. Who are the national heroes and what are their virtues? Besides political heroes, what other heroes are there (businessmen - literary men?) and what are their virtues? Are there also contemporary heroes and in what area or are there only heroes of the past (past-future orientation)? Are history and geography presented as the determinants of action or as the materials for action.

c) Priority of importance. Is it a humanistic-scientific curriculum; a vocational-general curriculum? Operational index: what are the consequences of flunking music or literature versus flunking math? Are there any required courses of programs? Which teacher has the highest academic prestige among his colleagues and which teacher among the pupils?

III. Method of teaching.

a) Subject centered - Child centered. Or, emphasis on knowledge versus emphasis on motivation. In the European countries, the teacher in the elite school has received a university degree in a particular subject-matter and he considers himself to be primarily an academic expert rather than an educator. In addition to the professional identification, a second determinant of the relative emphasis placed on subject-matter versus child is the position of youth and the authority relationships between the old and the new generation in a particular society. Several experiments have shown that the subject-centered, demanding teachers were found to elicit from their students hostility, apathy, and other signs of withdrawal, whereas accepting and supportive teachers decreased anxiety among their students and produced greater interaction and positive feelings among the students. As for the cognitive aspect, subject-centered classes seem to do slightly better on examinations.

See: Rasmussen, G.R. "An evaluation of a student-centered and instructor-centered method of conducting a graduate course in education," Journal of Educational Psychology, Vol. 47, 1956, pp. 449-461. (It is important to remember, however, that the examination was a multi-choice test. What would have been the result if the students were asked an essay type of examination?)

Under certain circumstances, child centered instruction may increase the ambiguity of the teaching situation, namely, where the group members do not have the skills to set and achieve goals.

See: Mc Keachie, W.J. "Student-centered versus instructor-centered instruction," Journal of Educational Psychology, Vol. 45, 1954, pp. 143-150.

b)Process-Product. Or, emphasis on the subject-matter as a means for developing certain general skills and aptitudes versus emphasis on the subject-matter as an end in itself. This bears on the question whether information is stressed or the acquisition of general skills and attitudes. A good index for this would be: type of examinations. Essay type or test of knowledge of facts.

c)Permissive - Authoritarian. Is discussion allowed or does the teacher lay down the law?

See: Mc Keachie, W.J. "Procedures and Techniques of Teaching: A Survey of Experimental Studies," in Sanford, N. (ed.) The American College, New York, 1962, pp. 320-327.

According to Mc Keachie, the lecture method is found to be most effective in cognitive learning; the discussion method is found to be most effective in non-cognitive learning. The famous Lewin, Lippitt and

White experiment seems to suggest identical result. See: Lippitt, R. and White, R.K. "An experimental study of leadership and group life in Swanson, G.E., Newcomb, T.M. and Hartley, E.L.(eds.): Readings in social psychology, New York, 1952, pp. 340-355. The authoritarian groups were highest in quantity of production, while the democratic groups were judged to be better in quality. The same tendency for authoritarian leadership to result in greater quantitative productivity and democratic leadership to result in higher morale has been noted in other studies. See: Adam, S. "Social Climate and Productivity in Small Military Groups," American Sociological Review, Vol. 19, 1954, pp. 421-425. Much research on the effect of different methods of teaching is inconclusive, because no specification is made in these studies as to the characteristics of the students involved in the experiments. Contrary findings may be due to the fact that the experiments are performed on a student population which differs in significant aspects from the group in the original experiment.

IV. Kinds of positive and negative sanctions used.

a)Grades versus non-grades. Although I know of no major study which tries systematically to measure the effects of grading versus non-grading we would expect that in graded classes there will be a higher cognitive output, but the interest or involvement in the course subject matter may be lower.

b)Group sanctions - Individual sanctions. This raises the problem of cooperative versus competitive groups. See: Deutsch, Morton "The Effects of Cooperation and Competition upon Group Process," Human Relations, Vol.2, (1949), pp. 129-152 and 199-231. In the cooperative groups, evaluation and grading of students in an introductory psychology course, was done by group. Each member received the same grade, the group mean, and all members of the best group were exempted from a term paper. In the competitive groups, each member was ranked according to his individual achievement within the group and the highest ranking individual was exempted from the paper. Deutsch found no evidence of superior output in any of the two groups, but as for the non-cognitive aspect of learning: the cooperative group showed more motivation and friendliness toward each other.

V. What is being rewarded or punished?

a)Social - Intellectual Behavior. What is the teacher's definition of the ideal student: the bright student or the social active student? French education is, for instance, purely intellectualistic and social considerations play no role in the evaluation of pupils. Many observers of French education agree that this creates a climate in which there is little room for citizenship-training. See, for instance, Hignette (bibliography).

b)Outer appearance versus inner qualities. We would hypothesize that the more emphasis is laid on the importance of good dress and manners, the less emphasis is laid on motivating the lower-class student toward social mobility. The emphasis on outer rather than inner qualities need not result only from a middle-class bias. It may be the effect of the need to maintain order and discipline. A too strong stress on discipline can lead to the neglect of those factors which from a motivational point of view are crucial like withdrawal, hypersensitivity, etc.

VI.Who is being taught?

a)Males - Females or both. Not much seems to be known about the effects of co-education on learning.

b)Age of student population. Younger children learn better when they can identify with the teacher, that is, when the teacher exercises expressive rather than instrumental leadership.

See: Parsons, Talcott. The School Class as a Social System.

c) Homogeneous or heterogeneous with regard to age, race, intelligence or religion. From a general sociological perspective we may expect that when people are grouped together on the basis of a common characteristic, this will increase the homogeneity among them and reduce the social distance between them. When, in addition, the characteristic which forms the basis for segregation carries different prestige (race, intelligence), the social distance from others may be translated in terms of social superiority and inferiority. Under these circumstances, homogeneous grouping will affect the motivation of the inferior group. Most studies do not show a difference in cognitive learning between homogeneous and heterogeneous ability grouping, except that the most intelligent children are slowed down somewhat in the latter.

See: Mc Keachie, o.c., pp. 337-338.

An excellent (but not frequently cited) piece of research was done by Atkinson and O'Connor. From their work on achievement motivation they hypothesized that students who were more strongly motivated to achieve than anxious about failure, would do better in an ability group class and would show enhanced interest in schoolwork. In contrast, it was hypothesized that those more strongly disposed to be anxious in a competitive achievement situation and less positively motivated to achieve would show less interest and satisfaction in schoolwork and would do less better in performance. Thus, ability grouping was expected to have differential motivating effects depending on the personality of the student. Two experimental studies proved to support the hypothesis.

See: Atkinson, John W. and O'Connor, Patricia. Effects of Ability Grouping in Schools related to Individual Differences in Achievement-related Motivation. Ann Arbor: Office of Research Administration, 1963.

(We would like to add that, possibly, the strength of the achievement-motivation and the anxiety about failure is related to social class background. If it can be shown that lower class upwardly mobile students have a strong anxiety about failure, then ability grouping in the school system makes it more difficult for them to succeed than for the secure, achievement-oriented, middle- or upper-class boy).

P.S. We should include under this heading also homogeneity or heterogeneity with respect to social class. The class composition of a school affects the aspiration level of its students.

See: Wilson, Allan B., "Residential Segregation of Social Classes and Aspirations of High School Boys," A.S.R., Vol. 24, 1959, pp. 836-845.

d)Peer group organization.

See: Gronlund, Norman E. Sociometry in the Classroom, New York, 1959.

The author does not specify what a good sociometric structure is, but he seems to suppose that when you have a class, where there is a high degree of interpersonal contact and where there are no cleavages, there will be a high academic performance and a strong motivation to do well. In view of the work of Coleman (The Adolescent Society) and in view of what we know about industrial sociology and informal groupings, this is, of course, too simplistic. Whether a good sociometric structure will lead to high academic performance depends on whether the class accepts the school values or not.

VII. Methods of Evaluation used.

a)Selection versus guidance. This distinction parallels the subject centered - child centered dichotomy. If there is a strong emphasis on subject-matter and knowledge per se, evaluation is restricted to a test of cognitive knowledge. Those who pass the test are promoted, the others fall behind or are selected out. Although there are no clear-cut data to rely upon, we may speculate that the more rigid the selection system, the lower the motivation of those who are selected out. The lack of attention to stimulating motivation may work primarily against the lower-class child whose motivation to perform to the best of its knowledge cannot be taken for granted. In the European educational system the evaluation of knowledge is used predominantly for exclusion or for selection rather than for diagnostic or guidance reasons.

b)Type of examination.

See: Bloom, Benjamin S. "Testing Cognitive Ability and Achievement" in Gage, N.L. (ed.), Handbook of Research on Teaching. Chicago, 1963, pp. 379-397.

Examinations as such affect knowledge, but also motivation to the extent to which doing well on examination is part of the academic reward structure. In addition, the type of examination used can influence the type of knowledge and aptitudes acquired. Bloom reviews three studies in which students were subjected to two different types of examinations: an objective test (multiple-choice) and an essay test. The students preparing for the first type of examination concentrated on the memorizing of facts, while the students preparing for the second type of examination tried to organize the material and to develop an own point of view about it.

VIII. Size of Class.

The effect of the size of a group has been studied by the small-group people.

See: Bales, R.F. and Borgatta, E.F., "Size of Group as a Factor in the Interaction Profile" in Hare, P. , Borgatta, E.F. and Bales, R.F.(eds.). Small Groups (New York, 1955), pp. 396-413. The greater the size, the greater the role-differentiation between leaders and followers. Size limits the possibility of involving the students individually and relying on their active participation. Motivation is likely to suffer in such a situation. This, in its turn, forces the teacher to stress the need of order and discipline in order to be able to control the non-motivated. It would be wrong to call -as is often done-all instances of the teachers' emphasis on order and discipline evidence of a middle-class bias. In large classes, this is simply an organizational necessity, required regardless of teachers' background.

IX. Scope of School Activities.

Is the school only an academic institution or does it also offer extra-curriculum activities? Participation in the latter may provide opportunities for citizenshiptraining and for creating a positive attitude with regard to the school. With the exception of the English schools, European schools are mostly academically oriented. Certainly at the vocational and the secondary modern school there are no school clubs. This means that there are no alternative channels of reward and motivation besides school success. On the other hand, too much emphasis on extracurricular activities may lead to an overemphasis on other than school values, as indicated by Coleman's study.

The relationship between the means-model and the goal-model is as follows:

	Ia	Ib	Ic1)	IIa	IIb	IIc	IIIa	IIIb	IIIc	IVa	IVb
Producer a)-g)				x							
" f)	x										
Citizen a)-e)				x							
" c)					x						
" d)									x		x
" e)									x		
Culture Consumer a)-b)				x							
" " b)						x					
Adult a)-b)				x							
" a)											
" b)					x						
General influence on:											
Knowledge			x				x	x	x	x	x
Aptitude								x	x		
Motivation		x					x		x	x	x

The following relationships among the variables in the means-model may be postulated:

A) The lower the social background of teachers, the lower their prestige. The lower the prestige of the teacher, the lower his perceived competence. The lower the prestige of the teacher, the greater the difficulty of maintaining order and, consequently, the greater the likelihood of authoritari teaching methods.

B) The greater the emphasis in teaching on subject-matter, the greater the emphasis on intellectual aspects in the evaluation of pupils, the greater the emphasis on selection as a way of evaluation and the narrower the scop of school activities.

C) The greater the size of the class, the greater the emphasis on authoritaria ways of teaching.

D) The greater the emphasis on groups rewards, the stronger the cohesion of t peer group.

	Va	Vb	VIa	VId	VIIa	VIIb	VIII	IX
roducer a)-g)								
" f)		x			x			
itizen a)-e)								
" c)								
" d)	x							
" e)			x				x	x
ulture Consumer a)-b)								
" " b)								
dult a)-b)								
" a)								
" b)								
eneral influence on:								
nowledge				x		x		
ptitude						x		
otivation			x	x	x		x	x

D. APPENDICES

3. Problems in the Theory and Method of the Classification of Outcomes

Anthony H. Smith

and

Jan S. Smith

PROBLEMS IN THE THEORY AND METHOD OF THE CLASSIFICATION OF OUTCOMES

Anthony H. Smith
Jan S. Smith

November 1967.

I THE CLASSIFICATION OF EDUCATIONAL GOALS

The first task we embarked upon was the classification and analysis of the statements collected by our fellow research assistant from the programmatic and philosophical literature on education. These statements, which had hitherto been referred to in this project as "educational goals" consisted of statements about education containing an evaluative element. That is, statements of what "should be" rather than what "is". Our major objective at this stage was the attempt to build inductively a classification scheme which would handle these statements of the desirable. In this task we first transferred all the statements of "educational goals" to index cards, and attempted to sort these cards into piles in which we believed the statements to share some common characteristics. We were, however, prevented from being able to establish such an inductive scheme by certain problems.

I, 1 Problems involved in attempting to build an inductive classification scheme

Three types of problems confronted us:

I, 1 (a) The use of the term "goal"
I, 1 (b) The definition of education
I, 1 (c) The meaning of specific goals as stated by their authors

I, 1 (a) The use of the term "goal"

As will be seen, a great deal hinges on what we mean by the term "goal". This problem is not merely one of definition. It is, rather, one of sifting through the various meanings of that term in the literature and trying to find some group of meanings internally consistent and clear enough to use in research. The first step in this procedure might logically be analysis of the theoretical ideas involved in the term goal. In the course of our research, however, it was collection of data (that referred to as "educational goals" above) which might serve as indications of educational goals.

It was hoped that inspection of these statements would aid in developing a conceptual scheme which could be used to classify educational goals. However, the diversity of statements made it apparent that a purely inductive approach would be highly inefficient. We needed a *a priori* conception of our universe of discourse, the things we were talking about, before finer distinctions would emerge from the statements themselves. Hence, the task was to find a clear meaning for the term "educational goals".

Two families of theoretical ideas surround past uses of the term goal in the sciences. One may be suggested by an approximate synonym, "subjective intention". We might separate out three presumptions on the part of theorists who hold that subjective intention is a highly useful concept. First, they assert that in many circumstances individual persons have definite conceptions of alternative, future states of affairs. Second, they assert that individuals have stable preferences for the realization of some states of affairs rather than others. Third, and crucial, they assert that these definite conceptions and stable preferences are determinately related to many other highly general aspects of behavior, so that they will serve as explanatory concepts in cases where our concrete observations of behavior appear to be unrelated to one another. If we are prepared to accept these presumptions, then we might employ the idea of goals as "subjective intentions".

The second general meaning of the term goal differs from the above in fundamental ways, and confusion of the two is therefore inimical to our work. This second meaning is applicable for any system of properties or variables, whereas the first was applicable only for individual persons. We say that a system of variables behaves purposively or has goals, if some variable in the system approaches or attains a specific value or magnitude -- called the system's goal --- regardless of environmental variations. Accordingly, we may look for as many goals of the system as there are variables in the system and we may discover that with respect to some variables the system is purposive, whereas with respect to others it is not.

The primary advantage of the first definition is that it is coordinated with the data. Subjective intentions are preferences of individuals; our statements can be understood to reflect personal preferences. The primary disadvantage of the first definition is that it does not allow us to speak about the goals _of_ educational systems, since educational systems cannot be construed as having subjective intentions. In fact, there is no way in which subjective intentions can be intrinsically tied to education, as contrasted with other institutions.

The second definition is valuable on this very point where the first definition falls short. It is applicable to any system of variables. Therefore, it allos us to talk about the goals of educational systems. But the second definition of what we mean by goal is deficient on two counts. First, it cannot be used in connection with the collected data on preferences of individuals. Second, in order to use it, we need advanced and fixed ideas of what educational systems constitute, and a large amount of aggregate data to identify the goals. These requirements have made the second definition of little use in social science, except in certain limited applications to economic organizations. But even in that case, there is much ambiguity and dispute about what the "goals" really are.

It appeared that we should take elements from both of these families of theoretical ideas to define the universe of discourse for this project. From the first idea, that of subjective intention, we take the idea that people prefer some conceived alternatives over others. We drop the requirement that these alternatives lie in the future. From the second conception, we take the implicit idea that educational systems may have ordered consequences. In combining these two elements, we focus on the preferences of individuals for educational systems. Hence, we may speak of the preferred outcomes of education.

I, 1 (b) The definition of education

This decision led us to another task in concept formation. It would be necessary to identify a set of variables called, collectively, education, and to isolate the effects of these variables on others, called outcomes of education. If this could be accomplished we could attend to the problem of interpreting the data, a problem which would involve the development of rules relating the statements of preferences to the logically possible outcomes of educational systems.

Two methods present themselves for identifying what we wish to subsume under the term education. The first is enumeration, that is, naming every element which we wish to put in the set of things called education. Enumeration did not recommend itself because we needed some abstract rule which would allow us to distinguish consistently among education, outcomes of education, and irrelevant data. Thus, it appeared that we would have to find some abstract definition of education. The ambiguity which resides in the term can be made clear if we try to distinguish between education and socialization.

Two distinctions were rejected because they appeared to restrict the study to an unduly narrow scope. We did not wish to consider education only as cognitive learning, thereby neglecting such phenomena as motivation and moral conformity. Nor did we wish to limit ourselves to socialization in organizations designed to socialize, i.e., the schools, since this would preclude assessment of the relative effectiveness of schools and other educational agencies. Other distinction, such as "education as institutionalized socialization", or, "education as socialization resulting from deliberation of the parties involved", were either ambiguous or over-abstract.

But if there are no restrictions imposed on the term education, if it is indeed synonomous with socialization, then by consensus of text books and theorists the outcomes of education are coterminous with the subject matter of sociology. And this is the rub: there are no viable conceptual schemes in sociology which are sufficient to apply to all learned behavior. At the least, such a scheme would be exhaustive and mutually exclusive at each level of abstraction, would be accompanied by rules relating category names to empirical observations, and would generate with the aid of a small number of premises, logically related statements of interdependence between variables. One attempt was made to begin the development of such a scheme, drawing upon materials developed by other theorists and simplifying them considerably. Two difficulties were inherent in this effort. One involved the relation between attributes of individuals and attributes of groups; the scheme developed was not capable of handling the latter. The second was the specification of rules which would coordinate the concepts in the scheme with observational procedures; no such rules could be developed without an effort considerably beyond the time constraints under which we worker.

I, 1 (c) The meaning of the specific educational goals as stated by their authors

Before reviewing the scheme which we eventually used, several problems raised by the goal-statements themselves, apart from commitment to any scheme for classifying them, can be noted. The most prominent and pervasive difficulty stems from the fact that the images of educational outcomes in the minds of individuals may not have any empirical content. For example, when one finds such statements as "education in America should provide each individual with the maximum creative freedom" or "our schools should teach the future generations to be effective democratic citizens" or "American education is the foundation of the laissez faire system", without any contextual statements to provide some sense of the author's intent, it is not possible to assign empirical content to the statement without great risks of error. It is impossible, in other words, to decide which of the possible outcomes of education the author prefers. And there is the further probability that he doesn't mean anything, semantically. Rather, his statement has pragmatic import; the symbols evoked -- symbols like individualism, creativity, free enterprise, equality, democracy, etc. -- have their meaning in the warm feelings and we-feeling people experience upon hearing them.

I will briefly touch on the evidence for considering a large number of purported goal-statements as elements in a quasi-empirical or non-empirical belief system. First, one could hardly expect less from educational institutions, since both economic and political institutions are widely recognized to state "goals" in a ritualistic manner, without hope of attaining them. In fact, Herbert Simon states the empirical generalization, the most abstract goals of the organization are never operational, which is to say that no one seriously considers achieving them. Secondly, many of the terms one finds in implicit descriptions of what the American school does to American society are patently inapplicable to

American society. Perhaps the best example is the one given above: "laissez faire." It is difficult to take "goal" espousals seriously from a person who describes the American economy as laissez faire. And, finally, there are many terms which are never used with consistent and specific empirical referents, terms like creative potential, and freedom, and individualism.

Note: For a further discussion of these problems, and one which throws some light on the process by which we came to realise the difficulties confronting us, the reader is referred to a brief paper that was written soon after we commenced the task of building an inductive classification scheme. This paper is included at the back of this report as Appendix One.

I, 2 The classification scheme developed at this time

Recognition of the limitations which would inevitably characterize a classification of the statements of preferred educational outcomes, as outlined above, was beneficial insofar as it allowed us to work without illusions about what the final product would achieve. We hoped to approximate the ideal conceptual scheme as closely as possible, and to achieve the minimum criteria:

1. classification of all statements taken from the literature on education
2. classification of each statement in one and only one category
3. reliability between coders in this classification

In spite of the above difficulties, the statements of preferences were eventually placed into a set of highly formalistic categories. This formalism appeared to be necessary in order to take into account the diverse conceptions of preference reflected by the statements. A large number of statements were eliminated as obviously without semantic meaning; in this class were such terms as maximum potential, best skills, etc.

The set of categories finally evolved were as follows below:

(a) Characteristics of Schools

- (i) Characteristics of Actors
    - Administrators
    - Teachers
    - Pupils
- (ii) Characteristics of Relationships

  (Any possible relationship involving any of the three types of actors).
- (iii) Characteristics of Facilities, e.g., buildings and books.

(b) Characteristics of Extra-School Phenomena

- (i) Psychological characteristics of actors outside of school roles
- (ii) Characteristics of relationships outside of the school
- (iii) Characteristics of the Society
    - Economic Aspect
    - Political
    - Familial
    - Etc.

Note that in the actual development of a scheme a primary distinction was between school and non-school, a distinction which had been rejected in policy discussions.

II Summarizing the Articles and Books

Having finished our examination of the "goal statements" or "preferred outcomes" as we would rather call them, we turned to our next task, that of collecting and summarizing existing theory and knowledge about actual educational outcomes.

II, 1 The problem of which literature to summarize

It was evident from the beginning that we would not be able to cover all the research and theory about educational outcomes which has been written in the last twenty-five years. If we had attempted this task, and in the highly systematic form in which we completed what summarizing we did do, our task would not only not be complete to this day, but would also need to be continued for several years. We had thus to be selective. This selection involved (a) the establishment of certain general principles so that we focused on some research rather than others, (b) the ruling out of certain specific areas which we know to be adequately summarized already and (c) finally, along the lines of (a) and (b), specific decisions over which literature we should take.

II, 1, (a) The establishment of general principles

Up to this time a relatively crude model of the (educational system" had been implicit in our thinking. This model was one of a process institution (schools) in which entering pupils were the "raw materials" for whom activities within the school system are assumed to have certain consequences. It is these consequences which have been labelled "outcomes" and in which we are interested. It is evident, furthermore, that the consequences of the school system for children are not wholly a function of that system, but also of prior-formed characteristics of the children themselves. Thus, for example we know that "the culturally disadvantaged child" performs less well in school than his "non-disadvantaged" counterpart; similarly a great deal of information exists about the consequences of stratification for education. It was thus at this point that the three-fold classification of variables which came at a later stage to be built into the card-file was arrived at. These categories are not to be treated as rigidly defined since as will be shown below in the discussion of how variables are classified. The category within which a given variable falls is frequently determined by its role in a theory. The categories themselves are:

(i) inputs (to the school system)
(ii) (school) system characteristics
(iii) outcomes (of the school system)

In (i) was included characteristics of pupils (broadly "home background") and factors in the local community (e.g. median income; avilable taxes) which are considered to shape the school system itself.

In (ii) was included for example characteristics of personnel and resources, plus any other possible conceptualization of variability within the school system. Also included here was "peer group" where this was construed as an unintended consequence of bringing pupils together within a school.

In (iii) was included any known or hypothesized consequence of children's sojourns within educational institutions. This category we quickly realized was unclear, a problem which will be discussed below.

First, however it should be noted, that given our present lack of knowledge concerning the limits of possible or even actual educational outcomes, research and theory about learning *outside* schools is potentially relevant. If, for example, we could conclude that behavior pattern *x* in school-pupils was consistently a consequence of family pattern *y*, then we could safely discount statements that suggested *x* to be a consequence of school pattern *z*. To include such research was totally unpracticable with a limited time period, and this research and theory (known conventionally as "socialization") could not be studied. Moreover, in the three categories above, our primary focus is on the relations between (ii) and (iii) with (i) to (ii) secondary.

But, to return to outcomes, it was quickly realized that our conception of outcomes was inadequate. Our original task was envisioned intuitively as including as "outcomes" anything which we could perceive as a consequence for pupils of immediately post-school adults of formal education. This would, in fact, have ruled out a sizeable body of knowledge concerning the relation between "recoverable" education variables (e.g. years of schooling; academic qualifications) and other variables (for example family patterns; prejudice; voting behavior) among adult populations. The reason that these types of research may be relevant centers on the fact that theory about these relations is either unproven, or non-existant. In other words, whilst we can establish various correlations between education and other variables amongst adult populations, we do not know "why" these relations exist. To answer this same sort of "causality" must be established. The establishment of such conventionally involves determining temporal priority between the variables (seldom in doubt is the case of education: in general people get educated before they get divorced) and specifying intervening variables. Unfortunately, "education" is related to stratification in two ways: first, it is related to parents' socio-economic position, and second it is related to respondents' socio-economic position. Theory about relationships between education and other variables with adult populations can thus invoke either:

(i) parental factors
(ii) the socio-economic variables complex
(iii) educational outcomes

This is best illustrated by an example. Supposing we know that: "The higher the educational level of any couple, the more likely they are to get divorced". This merely states a relationship and no explanation is evident. Any one of the three "meanings" of education specified above can be used to develop an answer to the "Why?" question. Thus:

(i) We can suggest that this is a pattern learned from parents which is perpetuated in children when they become married. This in fact involves shifting the relationship we have stated back in time and adding some assumptions about marriages taking place between people from similar socio-economic backgrounds, etc.

i.e. A. The higher the education of parents, the more likely they are to get divorced

B. The children of these parents marry into similar educational backgrounds

C. The greater the exposure of children to marital strife, the more likely they are to learn and repeat the behavior

D. The higher the educational level of parents, the higher the education of children.

From these statements it is possible to obtain our original proposition as applied to the "children" of the parents. In all these statements "socio-economic status" or even "income" could be substituted for "education" and the same relationships would hold.

(ii) We can develop an explanation directly from the fact that educational level is an approximation to socio-economic status. Thus, the differential divorce rate could be "explained away" by saying that marital strife is in fact constant at all socio-economic levels but that differential access to procedures for obtaining divorce (e.g. high cost) prevent lower income groups from obtaining divorces. This would require:

A. Divorce is expensive
B. Marital strife is unrelated to socio-economic status
C. The higher the income, the greater the possibility of paying the costs of divorce and obtaining divorce.
D. The higher the income, the higher the education

This again would give us the original proposition in which education acts as a measure of income which is the effective variable.

Finally, an explanation can be developed directly involving an educational outcome.

(iii) It could be maintained that the longer a child remains in school, the longer he is exposed to affectively neutral relationships and the less capable he becomes of maintaining affective relationships. This would involve:

A. School (as opposed to family and peer group) is the locus of the learning of affectively neutral relationships.
B. The more any individual is oriented to affectively neutral relationships, the less capable he is of maintaining affective relationships
C. Marriage is a predominantly affective relationship

From this again, it is fairly easy to see that our original proposition follows.

Unfortunately, from our perspective, it is not possible immediately from the original proposition to determine whether or not it is relevant for our purposes. We can, in fact, only declare that the relation does not involve an educational outcome where we can establish that either (i) or (ii) explains all the variance, or that (iii) explains none. Even where we can rule out (iii) we cannot establish for certain that there is not an alternative explanation involving a different educational outcome. For the reason, then, that data which relate "education" and other variables using samples drawn from adult populations may be extremely important sources of data about outcomes, we have included a sample of that research where "outcome" explanations might feasibly explain the findings. Thus, for example, we have included an article entitled "Marriage Patterns and Educational Level" (Glick, and Carter, American Sociological Review, Volume 23, Number 3, June 1958) in which only some explanations of relationships between marriage patterns and education are explicitly in terms of "outcomes" of education.

II, 1 (b) Ruling out of specific areas

Given the breadth of the focus described above it remained necessary to delimit the range of our material further. The next step was thus to rule out relevant areas which were already adequately summarized. Two major areas of research were felt to fall into this category.

(ii) "Equality of opportunity" research

This massive volume of findings and theory was not only our secondary sphere of concern ("input" research) but is also summarized in a number of places, particularly the "Robbins Report" in England, and the "Coleman Report" in the U.S.A.

Efforts to include both these types of research would not only have been time consuming, but also duplicatory. In addition to these areas we did not attempt to cover the reports which are published by the Office of Education. Indeed, in general we have tried to delve into literature in the field which is, is claimed to be, or can be viewed as, sociological in its approach, and which has not been summarized or critically examined by sociologists.

II, 2 The literature included

A list of sources for the material included in our file will be found in Appendix Two at the end of this paper. It should be noted here, however, that there are two major types of sources and that these two sources produce rather different materials. The sources are:

(i) sociological journals of the last decade
(ii) books in the "sociology of education"

The former tend to yield relatively complex and "sophisticated" theory and research, but focuses only on relatively crude measures of educational outcomes of the "years of schooling", "qualifications" or "academic performance" type. Data of this type are of course most readily available. The articles thus tend to be relatively unproductive in their yield of outcome variables, though time consuming to summarize because of their complexity. They do in fact tend to yield variables which are in the main "inputs" or school "system characteristics". They are, on the other hand, the most reliable and valid theory and research extant in the field and in our file.

The books with which we dealt, selected either because they purported to be sociology, or because they seem to have been generally influential in education, and have included (often unwittingly) sociological statements, yielded a much greater proportion of outcome variables per variable recorded. This differential productivity is explained by the fact that many of books are concerned entirely with speculating about educational outcomes, and do so without any systematic evidence, frequently without any evidence at all. It is thus relatively easy, and the more easy given that many of the books are directed at "educators" who do not criticize from the same perspective as sociologists, for the authors to speculate that "education leads to consequences A_1 A_2 A_3 --- A_n, B_1 B_2 B_3 --- B_n, etc. It is, on the other hand, much more difficult to develop a theory to explain why this is so, or to demonstrate that it is. Propositions which are taken from books are usually, therefore, simply of the worst sort. They have variables which are undefined and unmeasured, do not cite evidence and are in general "a-theoretical". We shall return to the meaning of some of these propositions when we come to discuss the way in which we have attempted to use the summary-file that has been constructed.

II, 3 The Techniques of Summary

Before we embarked upon an investigation of the research literature it was necessary that we devise a form for summarizing the literature which would be systematic and consistent. The technique we adopted was to take down "propositions" and "sets of propositions", definitions and measurement of the variables contained in these propositions (where definition and/or measurement occurred), and the evidence for the relations hypothesized. The term proposition refers to any statement of relationship between two or more variables, by which term is meant any "thing" which can be present in varying degrees. The simplest form of proposition is thus a statement of the relation between two "presence-absence" variables.

i.e. if A is present, then B will be present. Less simple propositions rely on more complex nominal scales or ordinal, intervals, or ratio scales. Interval and ratio scales allow us to state:

"The higher *x*, the higher *y*."

Propositions which contain more than two variables are more complex still. For example,

"The stronger the relation between performance and rewards, the stronger will be the motivation to perform."

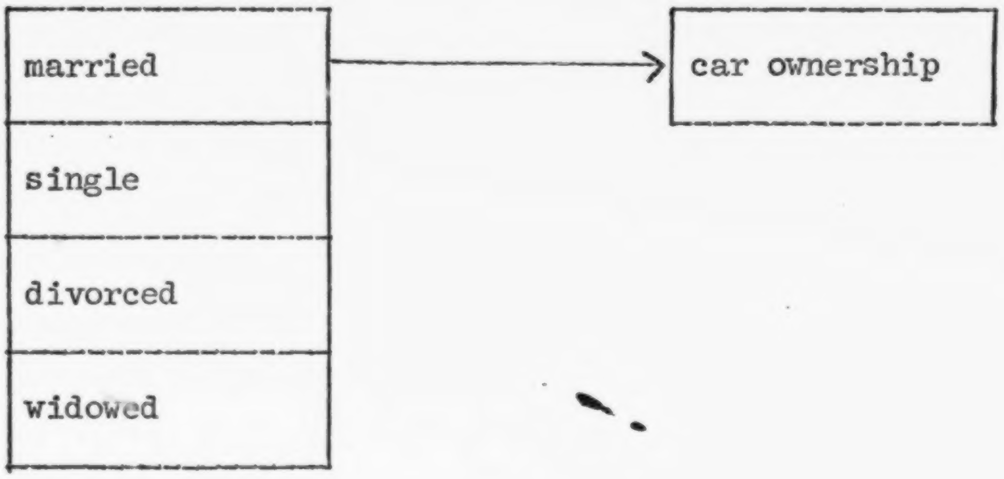
All propositions were taken down symbolically using "arrow" notations. Simple correlational statements are stated with double headed arrows.

$A \longleftrightarrow B$. A and B occur together.

Where direction (i.e. one variable determining the other) could be inferred a single headed arrow was used.

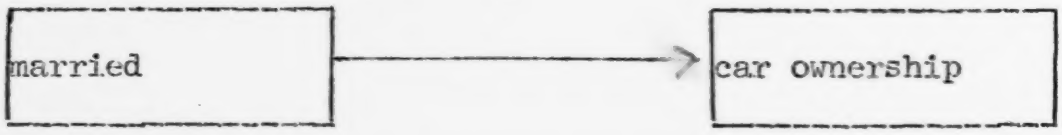
$A \longrightarrow B$. Where A occurs then B will occur. (This statement is more limiting since it does not allow us to infer "when B occurs then A will occur").

Nominal scales are the most difficult to express. Sometimes they are expressed in writing form, sometimes broken down into presence - absence variables. (It is often not clear in the sources whether certain variables are considered presence-absence, or nominal scales). A nominal scale is expressed as:



"If you are married you are more likely to own a car than single, divorced, or widowed persons".

or



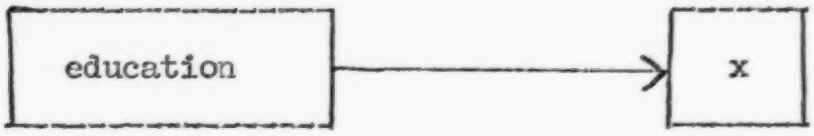
"If you are married you are more likely to own a car (than people who aren't married)".

With interval and ratio scales the words "higher" and "lower" and occasionally "stronger" and "weaker" are written above the arrows. Thus:

"The higher the education, the higher the income."

Examples of proposition sets can be found in the Appendix to this paper.

Any statement about an educational outcome falls within our concern with "propositions", since for any outcome x, the two variable statement:



is implied.

Once the propositions from any source had been recorded the variables were listed and definitions and techniques used to measure the variables taken down. In addition information which demonstrated any of the relations was also recorded.

The above is a description of the idealized form of the technique of summary, and the examples in the appendix are ideal examples. The propositional structure of the books was, as suggested above, much more sloppy than that of the articles and with the books we abandonned listing the variables since no definition, measurement, or evidence could be recorded. In its ideal form, however, we believe that the technique used to summarize the articles is the most explicit form of "note-taking" yet devised, and the form from which it is most easy to construct consensus over theory. The fact that we could not identify consensus is taken to be indicative of lack of such, and not to be traceable to the inadequacy of our methods of summary. (See below).

The actual task of summarizing involved considerable time expenditure on our part. Not only did we spend time developing and making consistent the form in which we were to summarize, but some of the literature we covered, in particular the articles in the "better" sociological journals, had extremely complex propositional structure. For example, Coleman's original article on the "Adolescent Subculture" took two and a half days on the part of one of us to reduce to a proposition set. This time was spent even though the researcher was familiar beforehand with the basic ideas expressed in the article.

Note: For a further discussion of the way in which material was summarized and an example thereof, the reader is referred to Appendix Three.

II, 4. An Assessment of the Summarized Material

The task of summarizing was, in part at least, abandonned for the reasons mentioned above; had we continued with it, as we quite easily could have done, we would have had no time to complete the tasks for which summary was but a preliminary. We can claim, therefore, only to have covered a portion of the literature concerned with education, and even only a portion of the literature which falls in the narrower focus set out above. Given the selected nature of our readings - the "best" research, the most influential books, books whose orientation was "sociological" - we cannot claim to be able to generalize these as "representative" of the whole field, since we have no statistical basis for so doing. We have, however, covered what we feel to be a near-complete range of the kind of statements which bear upon educational outcomes. This statement is, of course, made on the basis of intuition, and can only be defended on these grounds. At the time summarization was ceased, however, it was felt that no novel material was being dealt with in as much the literature was not yielding propositions which had not been encountered before.

III Constructing the File

The next task involved the attempt to systematize the summary that we had collected by indexing all the variables, and cross referencing them. Four stages were involved here.

III, 1 Writing each variable on a separate card and repeating this card for as many entries as were considered necessary for each variable.

This task was laborious and in the main merely clerical. Decisions had to be made concerning the number of entries to make in the file under any variable. These decisions were made on the basis of what were the most important words in any variable and the ways in which these variables were most likely to be approached by someone looking in the file. At first we erred on the side of a large number of entries for each variable: subsequently we realized that some of the entries were clearly redundant and cut down the number. At this stage we also had to decide how much information to include on each variable card. Time at this stage was pressing, and we decided to include the minimum of information, relying on references to the master cards to provide more. The following information was thus included:

(i) the variable

(ii) a designation of the "scope" or the variable, i.e. the answer to the question "To what person, group, institution, thing, etc. does this variable apply?"

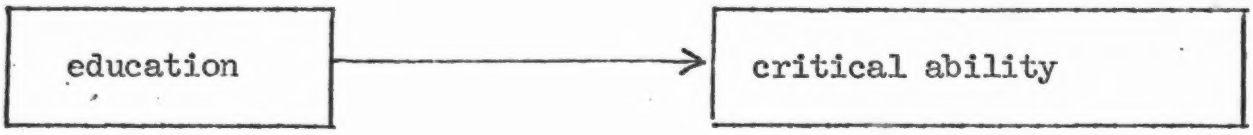
(iii) the source of the variable designated by name of author plus the number assigned to the particular source.

III, 2. Classifying the variables as "inputs", "system characteristics" and outcomes"

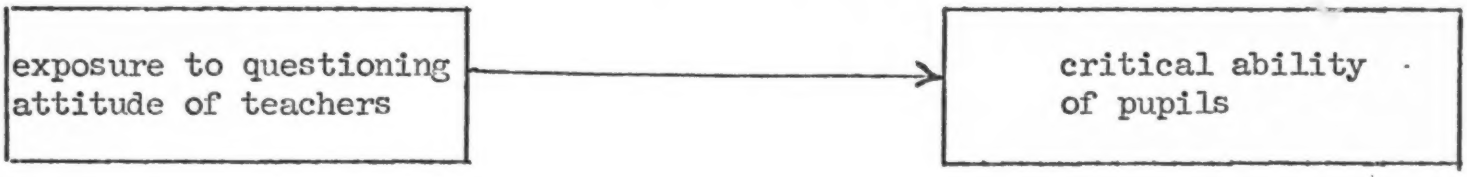
At this stage it was also necessary to classify the variables in order that the focus on outcomes be made more easy. The categories used in this classification are outlined above, and in this outline not only were the problems centering on the confusion engendered by the use of education as a stratification measure discussed, but it was also hinted that further problems inhere in distinguishing "outcomes" from "system characteristics". These problems again involve the way in which the variable "education" is used. Above three usages were delineated:

(a) as a measure of persistence of stratification-associated patterns over time
(b) as a measure of any socio-economic variable
(c) as a measure of some educational outcome

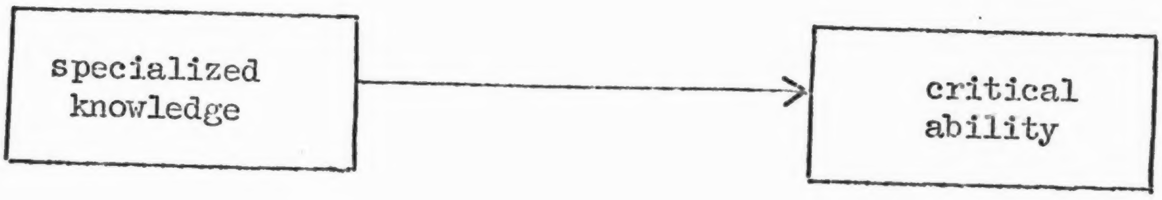
Besides these usages, however, it is possible to identify at least one other use of education. This is the use that is implied in the proposition:



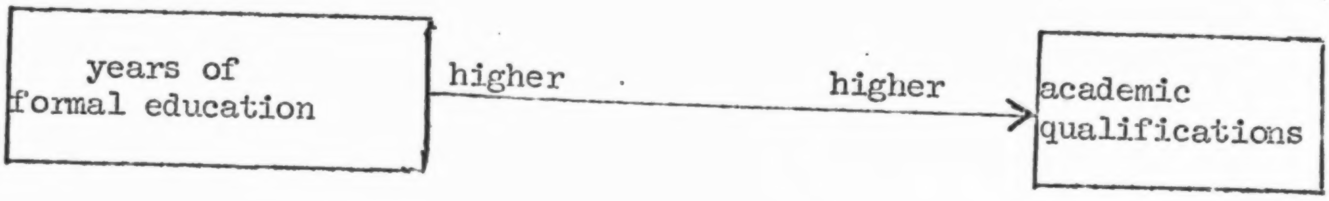
Here "education" can be construed as a measure of all or part of processes within the educational system. Reinterpreted as a more theoretical statement this might be taken to mean:


exposure to questioning attitude of teachers
critical ability of pupils

As with the example of education and stratification used above, the theoretical interpretation of this initial proposition could be completely other than this. For example it could be argued that specialized knowledge is necessary for "critical ability". Such that the proposition becomes:



Only in the latter interpretation is "education" a measure of an outcome. In fact "education" can be a measure of any school system characteristic, or any outcome, since anything in either category is obviously associated with say "years of schooling". This situation is further confused by the fact that problems of data collection lead to different measures of education being treated as interchangeable. For example many authors rely on the known relation



to use the two variables as substitutable. Thus even where the measure can be identified it is not always clear what it is a measure of, an outcome ~~as~~ or a system characteristic.

Given this argument we followed the same logic as we did in including the adult-population research. We are primarily interested in outcomes and therefore should not preclude any variable from being classified as an outcome which might conceivably be such. Thus, in general, education variables are either classified as outcomes, or double classified as outcomes and system characteristics.

III, 3. Alphabetizing the Variable Cards

This again is a clerical task the mechanics of which are self-evident.

III, 4. Reorganizing to ensure consistency and cross referencing

This proved a more complex task than had been envisioned. At the most simple level it merely involved deciding whether words were synonyms or not and then using a single word for several, or cross referencing where words were similar but were preserved as separate entries because of slight differences in meaning. At a more complex level, however, the problem at base of our two discussions of the "meaning" of the variable "education", the problem of the place of any variable in a more theoretical explanation recurred. Our hope had originally been to identify similarities in apparently different variables. In this we did not succeed. For example, the variable "intellectualism" might be considered by a researcher to be meassured operationally by the "number of books read per annum". This same measure might be construed by some other author as indicative of (i) literacy, (ii) aestheticism, (iii) knowledge, (iv) desire to increase own knowledge, (v) enquiring attitude towards world. Nor does this exhaust possible conceptualizations. It was not felt, however, that all these variables could, or should be treated as the same thing.

Ultimately the solution of such problems as these rests on (a) adequate general theory and (b) consensus over the operational meaning of theoretical terms. Since we cannot claim to have a scrap of either in the field being dealt with, we cannot reconceptualize at this complex level in any meaningful way. We further suspect that it is not feasible to approach the building of theory in the sociology of education by processes which are totally inductive. If a few similar theories existed, conceptualized within a single terminology, consistency would, we argue, have emerged from our file. That it did not is indicative of the confusion that reigns in theory about educational outcomes.

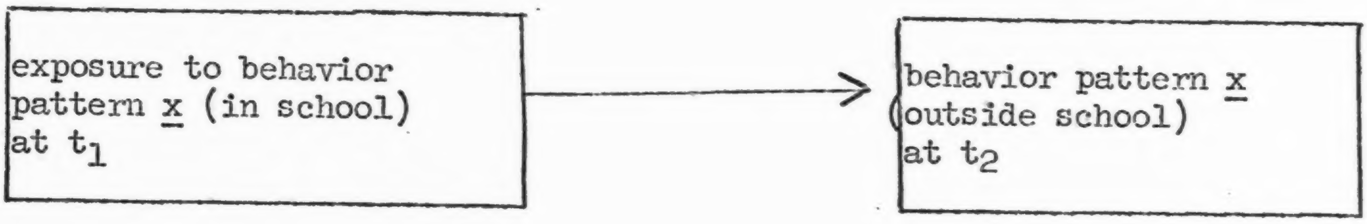
IV Attempting to Identify Consistent Theory from the File

The comments immediately above throw into doubt our next task, that of attempting to identify theory and/or findings about outcomes from the alphabetized file.

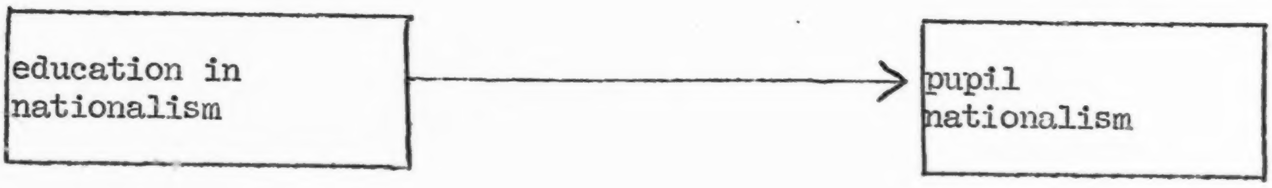
By "theory" about outcomes is meant some coherent and systematic statement of the ways in which variability in the organizational characteristics of the school system results in "pupil characteristics" or what we have termed here "outcomes". Such theory would ideally not only be consistent (i.e., various authors would specify the same outcomes as resultant from the same school-system characteristics) but it would also be deductively related. For example, the specific statements about the consequences of school system variability would be deducible from more general statements about the nature of learning and the most general processes involved in learning.

The technique used in the attempt to identify such theory was that of working through the alphabetized outcome-variable listings and referring to the master cards to see if the same variability within the school system was considered to have the same consequences by different authors. This task proved to be near-fruitless. Even by taking relatively large categories we discovered heterogeneous propositions. Two types of general proposition can be identified.

IV, 1. The first type is of the sort which locates the "learning" of a specific behavior pattern within educational institutions. The general form of this statement is:

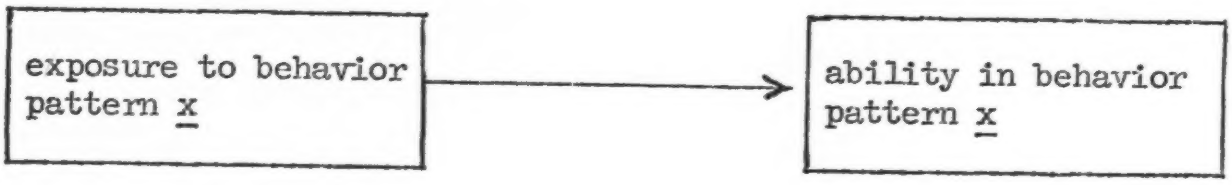

exposure to behavior pattern x (in school) at t_1
behavior pattern x (outside school) at t_2

A specific example would be:



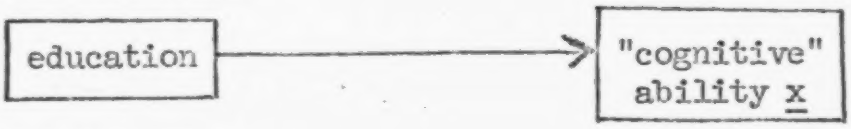
In the general form, the terms "teaching" and "education" can be excluded. It is difficult to define either of these by means other than the subjective dispositions of the teaching agents, i.e., by teachers' intentions that pupils learn. How and in what ways this "intent" influences learning itself remains to be shown (see Gage _op_. _cit_.).

Another form of the general statement occurs as:

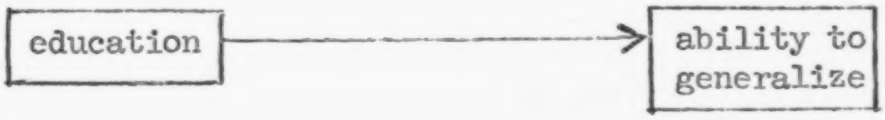


This type of statement is not meant to include those which have as dependent variable some analytic, psychological dimension. Rather it refers to statements which refer to a more or less specific behavior pattern from "machine tooler ability" to "occupational ability". Such statements appear to suggest merely that the behavior in question exists in memory in the time intervening between t_1 and t_2. Thus, in the general form of statments of the type cited above it would be possible to introject the "ability" variable as an intervening variable anywhere between t_1 and t_2. Indeed, dispositional states such as ability are generally inferred from two sources, (a) where there is evidence of the possibility of learning (i.e., from the first variable in the general statement) and (b) where the behavior in question occurs (i.e., from the second variable).

IV, 2. The second type of general statement which we can identify from the file is of the form:



The dependent variable here refers to conceptualizations of abilities which "cross-cut" specific action patterns; that is to analytic dimensions psychologically conceptualized. A more specific example would be:



The problems involved with such statements are evident from the fact that we cannot state (and the author's do not state) the independent variable at the level of specificity of the dependent variable. In fact, we do not know here what theoretical term education stands for and can only interpret education as say "years of schooling" or "level of specialized knowledge". That is, we do not know which of a range of system characteristics, both known and unknown, which may be related, quantitatively or qualitatively to the education variable, is actually the effective variable in producing "ability to generalize" in pupils. It is of course possible that the school system in all ~~has~~ its many and varied aspects accounts for the characteristic. This is a priori unlikely, and what is more, if it is true, it is even less likely that the school system results in the same way in other similarly-conceptualized variables.

In this context, the reader is referred again to the various previous discussions of the usages of the term "education".

IV, 3. Consistent Evidence

Next we must answer the question "What evidence bears upon the above propositions?" The answer is "very little", and what little there is is impressionistic, and will not stand assessments of reliability or validity. The first type of statement (IV, 1) is almost common sensical. Given that human behavior is learned, an assumption without which social science founders, it can be reasonably supposed that IV, 1 holds true and that there is some direct relation between length of exposure and amount learned. What the authors of such statements do not consider, however, is "forgetting". Memory is extraordinary fallible, yet there is no research which we have located which examines "retention" over relatively long time spans, or which attempts to specify which types of learned behavior are remembered, and which most easily forgotten. Besides this, moreover, it is possible to raise a second objection to such statements. Most of the authors seem extraordinarily naive in assuming that the teaching of behavior to children automatically results in that behavior when they grow up without any consideration of the adult motivational bases for the performance of that behavior.

Whilst it can be reasonably argued that the learning of a behavior is necessary before that behavior can be performed, it is unlikely that learning is a "sufficient" condition for adult performance. Thus, more concretely, the teaching of "Indonesia nationalism" in Indonesia's schools may not result in adult nationalism without, say, nationalist political parties which provide the opportunity for adult participation (see Fischer in Coleman, J.S. "Education and Political Development").

The second type of general proposition (IV, 2) is probably more easily testable since both variables are in principle measureable using existing techniques (length of schooling; psychological testing). Standard intelligence tests appear to measure dimensions as "ability to infer", "ability to generalize", etc. Interestingly enough, however, these tests are deliberately constructed so as to deny the possibility of using them to measure increasing cognitive abilities. By introducing an age factor into the calculation of IQ scores the tests assume that these abilities remain relatively constant and users tend to attribute variations of individual's scores on the tests over time to inadequacies in the tests or motivational factors, and not to "intellectual growth". The usefulness of these tests for factoring out ability is, of course, obvious, but they are of least use for examining the range of outcomes conceptualized as generalized psychological "abilities". We thus have no discriptive range of data even on these outcomes.

The two types of statement (IV, 1, and IV, 2) together with evidence constitute virtually all the summary we can make of the file. Any further summary would merely involve taking more specific statements of either general type, or taking other unique propositions and listing them. This is a disturbing conclusion in view of the fact that we had hoped to be able to build, if not consistent findings, then at least consistent theory out of the file. Why can we not do this? There are of course two possible answers:

IV, 3, (a) The file itself is inadequate
IV, 3, (b) No such theory exists, whether implicit, or explicit

IV, 3, (a) We cannot of course rule out such an explanation. The file is less than ideally comprehensive, and the literature on education vast. Moreover, we have deliberately neglected two areas (although we can safely assume that the area of these two most relevant to our concern with "outcomes", that on "teacher effectiveness" has yielded as yet, no reliable findings). The file was, further constructed in a relatively short time, and a time during which we had to solve all the problems connected with its construction. Any one, or all of the problems listed above may detract from any conclusions we draw from the file: for example, a cogent and comprehensive theory concerning educational outcomes may exist in the file obscured by the fact that we have been unable to get consistent conceptualization of the variables. Certain things, however, mitigate against the denigration of both the file and our conclusions from it:

IV, 3, (a) (i) We are not only reliant on the file, but also on our broad impressions of both the literature in the file and other literature in the field of "sociology of education". In our reading, we have not at any stage been aware, even on any intuitive level, of any emerging theory beyond the two statements posed above.

IV, 3, (a) (ii) The inadequate propositional form of most of the variables is indicative of the fact that they are not part of a general theory. It was noted above that the books yeilded a higher proportion of outcome variables, and a higher absolute number of such outcomes than did the articles. The books, however, defined the terms less well, gave no evidence and usually produced propositions which were comprised of two isolated variables. If the statement:$x \longrightarrow y$ occurs, we are immediately prompted to ask why these variables should be related. If an answer is supplied it usually emerges as a "proposition set", i.e., as a series of interconnected variables. (See for example the expansions of the proposition about education and divorce above). The absence of such sets from the book sources is thus indicative of the fact that the authors do not answer our "why" questions and perhaps cannot answer these questions given the paucity of theory elsewhere in the field.

IV, 3, (a) (iii) In dealing with the goal statements we realized that the authors of these statements operate within a framework which allows them to attribute any action, or any abstraction from action to the educational system as a goal. The same applies to the statements about outcomes. It is not immediately apparent either to us, or to the theorizers about outcomes, why any behavior pattern cannot be learned within a formal educational institution. The range of outcomes in the file thus seems to cover "any action". At the same time, it is unlikely that this is the case, since the organizational characteristics of schools are likely to set limits to the variability of what they can teach, limits which are considerably narrower than "any action". Only two articles with which we are familiar actually attempt to approach the problem by taking organizational characteristics of the school, contrasting these with characteristics of family and peer group, and delimiting likely outcomes. This is probably the most worthwhile approacl to the problem. (Hence one of our early papers tried to use a similar approach) The two articles are:

Parsons, Talcott "The School as a Social System".

Because it was written over a decade ago this does not appear in our file. And:

Dreeben, Robert "The Contribution of Schooling to the Learning of Norms", Harvard Educational Review, Vol. 37, No. 2, Spring 1967.

IV, 3, (b) Some Conclusions about Theory

These considerations (IV, 3 (a) (i), (ii) and (iii)) lead us inevitably to our second explanation for not having been able to identify theory, viz. that no adequate theory about outcomes exists, beyond the general propositions identified and the two specific articles mentioned above.

V THE CLASSIFICATION OF OUTCOMES

Having established very little consistency from the file, and because we were increasingly pressured for time, we turned next to the attempt to classify the outcome variables in terms of the scheme developed for this purpose by Professors Tumin and Bressler. It was hoped that by using this scheme an alternative way of generating some cogent and systematic theory might be found. In this section of the paper, we will present a short description of the scheme and the procedures used in classifying the variables, some examples of classified variables which may help to clarify the meanings to be associated with the various categories in the scheme, and a summary of the main problems involved in using the scheme.

The scheme was designed to take into account variations in "outcomes" that could be attributed to variable characteristics within the school system. It was designed to classify pupil outcomes (e.g. "creativity", "literacy", "capacity for critical thought") and not consequences for the society(e.g. "industrial prosperity", "democratic electoral system", and "collectivism"). The link between the "micro" and "macro" levels is at present so poorly established in sociological theory as to preclude any pretense of valid inference from one system to another. The model, therefore, contains no provision for recording the impact of individual behavior on society.

The complete classification scheme can be found in Appendix Five where it is given geometric representation. Since it has three dimensions it can be represented as a solid, such that each outcome variable can (ideally) be placed in one of its cubic spaces. Each of these cubic spaces then represents the intersection of one category from each of the three dimensions. For example, we classify "educational aspirations" in:

Dimension	Category
(i) Domains	Activities involving cultural heritage
(ii) Facet	Knowing
(iii) Aspect	Psychic resources of roles

A short preliminary description of the dimensions may aid in understanding how they are used. The first dimension which we have called "Domains" locates the institutional sphere, or "part of society" in which the outcome occurs. The second dimension called "Facets of Domains" is used to analyze the behavior along the "psychological" dimension - "Knowing, Valuing, Doing". Third, the dimension "Aspects of Domains" is used to classify the outcome variables in terms of aspects of roles, hence the four most general categories of this dimension, "Requirements of Role", "Standards and Norms", "Specifications of Role , and "Rewards of Role".

As suggested above, the model is designed to handle "outcome variables". A discussion of the term "variable" will be found above in the passage dealing with the techniques used for summarizing the literature. In brief, by "outcome variable" is meant any characteristic of a person which is assumed to have been learned within school or influenced by school experiences. These variables are specified at fairly high levels of generality, and like all abstractions, denote differences in content, direction, magnitude and intensity. Thus the scheme permits us to record "level of educational achievement" and "capacity for affective relationships" in different categories, but neither is further differentiated according to "good" or "bad", "large" or "small", "high" or "low" and so on. The same applies to a qualitative attribute such as "marital status" which is not subject to quantitative distinctions. The designations "single", "married", "divorced", "widowed", all fall within the same category on the classification scheme.

The procedures used in classifying the variables were as follows. Once the variable had been identified by the procedures described above, and recorded as specified, we took each individual outcome-variable card from the indexed file, and assigned each of these to categories on the three dimensions of the scheme. This classification was recorded on each variable card. When we had assigned categories to about half of the variables, we returned to the beginning of the file and checked the code-recode reliability of our assignments.

Some examples of variables in the various categories should further illuminate the meanings associated with the scheme. In reading these examples, it may be helpful to refer to the appendix (five) in order to keep the whole conceptual scheme in mind. First we can illustrate the distinction between the categories in the Facets dimension. This can be accomplished most easily if we "hold constant" the other two dimensions, and look at variation only in the Facets dimension. Consider the following examples:

DOMAIN: Cultural Heritage

ASPECT: Requirements of Role, Knowledge

FACET AND VARIABLES:

(i) Knowing: "Academic achievement (Grade Point Average)"

(ii) Valuing: "Valuation on academic achievement"

(iii) Doing: "Application of knowledge to new contexts"

It is the Facets dimension which records, in this case, the relevant difference between these three variables. The other two dimensions record the fact that they also have relevant things in common, namely that they are all primarily involved in the Cultural Heritage part of society, and that they are all interpretable as knowledge required to perform roles.

If we look at the differences between these variables, we can get a better fix on the import of the Facets dimension. The most obvious distinction is probably that between the variable "valuation of academic achievement" and the other two variables. The former is placed in the valuing category for the obvious reason that it alone contains an element of "valuing". It alone gives us information about what a person likes and dislikes. The second distinction, that between "academic achievement (Grade Point Average)" and "application of knowledge to new contexts", hinges on how we interpret the Doing category. The variable "academic achievement" conventionally refers to the student's demonstration on some test or another that he has learned a certain amount of the subject of the subject matter taught within the educational system, that is, that he possesses knowledge of subject matter. Therefore, we put this variable "academic achievement" in the Knowing category of the Facets dimension. In the case of "application of knowledge to new contexts" we are faced with a variable in which knowing (knowledge") is made secondary to an activity ("application to new contexts"), i.e. knowledge itself is here an instrument in the execution of an activity outside the school system. For this reason the variable is placed in the Doing category.

The decisions in classification described above may be easier to understand if they are formulated as abstract "coding rules".

a) If a variable contains some reference to what a person likes or dislikes then it will be classified under the Valuing category.

b) If a variable does not contain a reference to liking or disliking, but denotes some measure of a memorized or learned activity not manifest as behavior, then it will be classified as Knowing.

c) If a variable does not contain a reference to liking or disliking, and does not denote some summary measure of knowledge, but denotes an activity (which may or may not require knowledge), then the variable is classified Doing.

These coding rules, however, do not exhaust the differences between the categories of the facets dimension. Consider these further examples.

DOMAIN: Political

ASPECT: Standards and Norms (Moral and Legal)

FACETS AND VARIABLES

(i) Doing: "extremist politics"

(ii) Knowing: "recognition of government as source of values"

If we look at the difference between these two variables we see that they are not explained by the above coding rules. Consider, first of all, "recognition of government as the source of values", which is certainly not a summary measure of knowledge learned within the schools. However, it must go in the Knowing category if we are to retain the conventional meaning of the words we use. Recognition is an element of Knowing, not of Valuing or Doing. We might say that recognition and knowing have the "cognitive aspect" in common. Furthermore, "extremist politics" does not really fit into the rules we have developed for assigning variables to the Doing category. In fact, we have put "extremist politics" in the Doing category because we could not put it in elsewhere. Hence we have further coding principles.

d) If the coding principles listed above do not work, and if a variable obviously contains a cognitive element, then put it in the Knowing category.

e) If the coding principles listed above do not work, and if a variable contains neither cognitive elements nor elements of liking or disliking then put it in the Doing category.

According to the second coding rule here, Doing is a residual category in the Facets dimension. Because of this fact, the Doing category contains a heterogeneous set of variables, only some of which are capable of being grouped under a single positive definition.

Moving on to the Domains dimension, we can again get a more definite idea of the coding scheme if we review several actual cases. Again we "hold constant" the other dimensions so that all relevant differences in our example are to be explained in terms of the Domains dimension.

FACET: Doing

ASPECT: Specifications of Role-Routines

DOMAIN AND VARIABLE:

1. Economic: "full time employment at time of high school graduation"
2. Political: "civic participation in community projects"

3. Primary Group: "marital status"
4. Cultural Heritage: "creativity of research techniques"

Another set of examples, from different categories of the Facet and Aspect dimensions, follow:

FACET: Valuing

ASPECT: Moral and Legal Standards and Norms

DOMAIN AND VARIABLE:

1. Economic: "acceptance of work discipline"
2. Political: "liberalism of civic attitudes"
3. Socialization/Education: "approval of school integration"
4. Cultural Heritage: "change [over time] in importance freshmen attach to grades"

We have said above that when we speak of a Domain we have in mind a part of society. Our use of the Domains categories in this manner must be made explicit because there is so often confusion about whether terms like political and economic refer to aspects of behavior or to major institutions in, or parts of society.

If we had to use the Domains to refer to "aspects" of behavior we would, for example have used "economic" to refer to "those activities concerned with the allocation of goods and services" thus including in the economic category parts of government organization which are here called "political". Similarly a definition of "Political" in the same terms (aspects of behavior) would have involved "those activities concerned with the allocation of power", a definition which would have included any variable related to the allocation of power in factories (classified under the scheme actually adopted by us as "economic"), or the allocation of power in any of the other institutional spheres.

Our use of "Domains" as parts of society may be more explicit if we list several words which jointly constitute near-synonyms for the respective categories:

Political: government, political parties, pressure groups, etc.

Economic: farms, factories, business firms, etc.

Primary Group: families and friendship groups

Socialization: schools, summer camps, universities, etc.

Cultural Heritage: churches, religious organizations, organized recreational activities, suborganizations of government, families, schools which are involved predominantly in maintenance of cultural heritage, etc.

These synonyms are neither exhaustive, nor perfectly clear, which reflects the fact that we have not found any method which allows consistent separation of the Domains in the coding process.

The category in the Domains dimension labelled Orientations to

Self and Others is residual, that is, it includes all or the variables which could not be assigned to one of the other categories. However, it does appear to include several sets of variables which have distinctive characteristics in common, and are therefore susceptible to positive definition. First of all, there is a set of variables which refer to "purely psychological" processes, without any specification of the social context in which they occur. Some examples are:

"tolerant attitudes toward others"
"aggressive behavior"
"happiness and well-being"

A second set of variables in this category refer to a dimension of social relationship which may occur in more than one of the five concrete domains:

"universalism-particularism"
"cooperation"

A third set of variables in the category all characterize relations between nonspecialized, "solidary" groupings (ethnic, racial religious, class) in societies. The generic example is prejudice toward the out-group.

The third dimension, Aspects of Domains, was definitely the most difficult to use. We have attempted to isolate the reasons for this difficulty, in order to present them here, though we are not certain that this analysis is correct. One problem was that we could not comfortably classify some variables in any category of this dimension; it does not appear to exhaust all of the outcomes which were taken from the research literature. Thus, the assignment of some variables to a category on this dimension is "forced". This problem had been solved for the other two dimensions because on each of these there was one category we could treat as a residual. With the Aspects dimension there was no obvious residual category.

A much more serious problem, however, was that the conceptual framework which implicitly underlies the Aspects dimension is more complicated than are most of the conceptualizations of the variables the were coded. The conceptual framework of the Aspects dimension it taken from the sociological analysis of institutionalized behavior patterns. Minimally, this framework involved (1) specification of a situation in which behavior is expected; (2) specification of who expects the behavior; (3) specification of from whom the behavior is expected; (4) specification of what sanctions are contingent on conformity or deviance from expectations. The same variable may fall into more than one of the categories in the Aspects dimension because it does not contain, implicitly or explicitly, information which allows us to place it within this conceptual framework (1 to 4).

In our initial investigation of the problems underlying the "Aspects" dimension we focused attention on its claim to be a classification of "Role-related" sets of activities. A role is conventionally construed as a relatively homogeneous set of expectations held by a number of persons of some "position" - thus "teacher", "doctor", "father", "pupil", "politician", etc. The actual homogeneity of the expectations is an empirical matter. As it turns out, however, most outcomes reported in the research literature cannot be interpreted in these terms. For example such variables as "abstract attitudes in political sphere", "ability to rationalize", "effectiveness of problem solving" simply have no readily discernable role referent. Even, however, when outcomes are conceptualized in role terms, these designations invariably assist us to locate their "Domains" but not their "Aspects". So, for example, "occupational skills", "citizenship skills," "skills of motherhood", and "teaching skills" refer to diverse institutional spheres ""Domains" in our terms" but only the word "skills" has any relevance for classification on the "Aspects" dimension. Accordingly the headings of the major sections of the Aspects dimension("Requirements of Role", "Specifications of Role", "Rewards of Role") are misleading in as much as any relation, any time allocation, any reward can be subsumed under these headings.

To what, then, do these sections of the aspects dimension refer. Speaking very generally they seem to describe:

1. Orientations to self and others.

   This category, of course, recurs from the Domains dimension, a problem the consequences of which we will discuss below.

2. "Psychological Characteristics" or "Dispositional States"

   This category refers to "capacities" and "motivations" to perform action patterns without these action patterns actually being present, i.e. actually specified as the outcome variable. The overlap of this with the first two categories - "Knowing"and "valuing" - on the Facets dimension is again obvious and will be further discussed below.

3. Standards and Norms

The whole subsection is here transferred from the original scheme with the addition of one residual category (see the expanded scheme immediately below). This category then refers to action either normatively (or ideally) defined, or action not so defined which falls in the added residual.

4. Culture

This category is what is left when action, and actors have been removed, i.e. it refers to "artifacts" and statements about the physical environment.

By subsuming the original scheme under these headings we are able partially to resolve ambiguities of classification which were otherwise beyond solution. In the following rearrangement the original scheme is set out on the left hand side of the page. The minor headings of this scheme recur in the new scheme on the right and have been retained for two reasons: first, to enable the reader to see how they are rearranged and second, since they are the terms around which our actual coding decisions were made. The letter and number designations of the subsections of the new scheme (right hand-side) refer to the location of these same sections on the old scheme and therefore do not appear in numerical or alphabetical order.

ORIGINAL

A. Requirements of Role

- (i) Relational
  - (a) Interpersonal
  - (b) Intergroup
- (ii) Knowledge
- (iii) Skills
  - (a) Mental
  - (b) Physical
- (iv) Resources
  - (a) Material
  - (b) Human
  - (c) Psychic

B. Standards and Norms

- (i) Moral and Legal
- (ii) Aesthetic
- (iii) Style and Manners

C. Specifications of Role

- (i) Routines
- (ii) Time Allocations
- (iii) Space and Movement

D. Rewards of Role

- (i) Property
- (ii) Power
- (iii) Psychic
  - (a) Interpersonal
  - (b) Intrapersonal

REORGANIZED SCHEME

1. Orientations to Self and Others

* to Self

D(iii) (a) Psychic - Interpersonal

D(iii) (b) Psychic - Intrapersonal

** to Others

A(i) (a) Relational - Interpersonal
A(i) (b) Relational - Intergroup
D(ii) Power

2. Psychological Characteristics (Dispositional States)

A(ii) Knowledge
A(iii) (a) Skills - Mental
A(iii) (b) Skills - Physical
A(iv) (b) Resources - Human
A(iv) (c) Resources - Psychic

3. Standards and Norms

B(i) Moral and Legal
B(ii) Aesthetic
B(iii) Style and Manners
C(i) Routines

4. Culture

A(iv) (a) Resources - Material
C(ii) Time Allocations
C(iii) Space and Movement
D(i) Property

This regrouping of the categories helps to explain the way in which we used the categories, since it groups together what we took to be alternative, or near-alternative categories. Thus, for example, the categories "*mental skill*" (A(iii)(a)) and "*psychic resource*" (A(iv)(c)) often occurred to us as alternatives for variables such as "creativity". Similarly, the categories "*material resources*" (A(iv)(a)) and "*property*" (D(i)) tend to refer to similar things, e.g. money and buildings.

The way in which this revised Aspects scheme was used implicitly by us in classification is complex since it was used *for those variables with which we could not immediately associate a category of the old scheme*. Thus, for example, whereas the variable "tendency to generalize affective relationships" is classifiable easily on the old scheme since it contains the term "relationship", near identical with the word "A(i) relational", the variable "racial prejudice" requires that we first ask, "What is this variable?" Our answer is that it is an "orientation to others", the "others" being a racial group, and *therefore* that it refers to a "*relation*" with a group racially defined.

The utility of rearranging the subcategories of the old scheme, however, can be demonstrated in another way; by showing how this rearrangement enables us to predict the distribution of outcome variables on the other dimensions of the original classification scheme. If the regrouping holds good, a number of "overlaps" with the other two dimensions of the classification scheme - Facets and Domains - are made clear. First, section 1 of the revised Aspects dimension is identical with section 6 of the Domains dimension: both are labelled "Orientations to Self and Others". If these two categories had been used to mean exactly the same thing, then it would follow that any variable classified under one of these sections would automatically fall in the other. This is not so for two reasons. First, category 6 of the Domains dimension was used residually so that not only does it include more variables than does its counterpart on the Aspects dimension, but even variables which are clearly "Orientations to Self and Others" only fall in this category if they do not contain a specific reference to the first five Domains. Nevertheless, from the overlap of the two categories it could be expected that most of the variables from "Orientations to Self and Others" on the Aspects dimension would fall in "Orientations to Self and Others" on the Domains dimension, and this turns out to be so. A total of *54 variables is classified in this way on the Aspects dimension and of these 36 fall into category 6; Orientations to Self and Others on the Domains.

Second, a similar prediction can be made for those sub-headings which are relocated under the revised heading "Psychological Characteristics". Since these refer to dispositional states rather than action, one would not expect variables located here on the Aspects dimension to fall under Doing on the Facets dimension. This picture ia again confused by our use of Doing on the Facets dimension as the residual for Knowing and Valuing, nevertheless, of a total of 226 Psychological Characteristics, only 15 are classified under the Facet Doing.

With the third category on the revised scheme, since "norms" imply valuing it could be expected that any variable falling under section (B) on the original scheme, or under the first three categories of (3) on the revised scheme would fall under Valuing on the Facets

---

* Note: See Appendix Five for the distribution of outcome variables over the classification scheme.

dimension. This is true for 33 out of 45 variables. Of the remaining 12, 8 fall into Knowing, an anomaly explained by the fact that it is possible to "know about valuing". With the residual category for standards and norms, "Routines" (C(i)) on the original, the category in which we place variables referring to action where "valuing" is not implied, we can predict that anything falling in this category will come under the Facet "Doing" which is used in exactly the same way for the Facets Dimension. This prediction holds true for 16 cases with one single seventeenth exception.

Finally, with the fourth category we cannot make predictions about the scatter of variables on the other two dimensions. One would expect not to have anything from the category "Culture" on the revised scheme in the category Doing on the Facets dimension. How does one perform something which is essentially a non-action? This would be so, had we not used Doing as the residual for the Facets dimension and placed in it any variable where we could not inter "Knowing" or "Valuing" so that for example 11 "time allocations" are "Done".

The problems involved with the Aspects dimension are thus considerable and suggest that this category above all needs extensive revision in such a way as to make it independent of the other two dimensions.

GOALS-MEANS ANALYSIS

1. For the moment we have abandonned the distinction between goals and means as originally used, since, as has been pointed out, it is possible to distinguish means-goals chains in endless sequences such that the position of a given preference statement relative to one other related statement, rather than the intrinsic characteristics of the preference statement, determine whether it is classified "goals" or "means". Thus for example the preference statement concerning

> "full utilization of intellectual powers"
>
> appears in M.11 as a means to the goal,
>
> "surviving as a democracy in an age of enormous technological and social complexity."
>
> whereas an almost identical statement to the former appears also as a goal (G.12)
>
> "helping each student achieve his optimum intellectual development."

2. We are thus, at present, dealing merely with preference statements: these appear to fall into two broad categories which in turn can be broken down.

   A) Preference statements which refer to desired characteristics of the school system.

   B) Preference statements which refer to desired "outcomes" of the educational process.

A) The school system

It should first be pointed out that it may be necessary to break this category down and introduce a third major category to include desired relationships between educational systems and other systems in terms of "inputs". This category would not only include the "equality of opportunity" debate, but also statement about who should control educational systems, desirable amounts of finance, etc.

Setting aside this problem, however, we now can see the preference statements about educational systems along the lines of the general "models" of those systems discussed at previous meetings. The preference statements can be classified as follows:-

(i) characteristics of actors

- a) pupils
- b) teachers
- c) administrators

We have yet to find an example of (c), though some may be expected to occur, as may more explicit desired role-occupant characteristics, e.g., desired attributes of "superintendants".

(ii) characteristics of relationships

These assumedly involve any relationship which is possible under (i) characterized in terms which need not concern us at present.

(iii) characteristics of "facilities"

(a) e.g., written knowledge

(b) e.g., buildings

These bear a remarkable resemblance (and for reasons which we shall come to) to the three general systems of "action theory" -- "personality system", "social system" and the residual "cultural system".

(N.B. Many of the statements which can be classified under this heading appear in the original classification scheme as "means" statements)

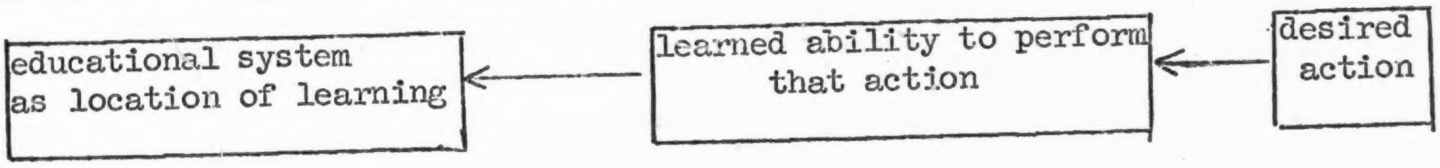
B) Outcomes

All statements classifiable as outcomes appear to refer ultimately to desired characteristics of the products of the educational system, that is to desired characteristics of erstwhile pupils. The form in which these statements may be made, however, is various. Thus, for example, the preferred outcome

"skills to the economy"

can only refer to actors possessing skills relevant to certain occupations, although the original appears to state a relationship between "educational system" and "economic system". This is merely another way of saying that in a "relationship" between two structures which is mediated only by the passage of actors from one structure to the other, then it is the intervening psychological variables (i.e. "characteristics of actors") which are the appropriate subject of study.

From this it can in fact be seen that in general we are dealing under this heading with statements of the following form:-



The logic of this process fairly obviously proceeds from right to left: that is, the preferred outcome is a desired action which is then attributed to the actor as a learned characteristic, and the learning process located in the school system. Indeed, both sociological theory and the range of goal statements which have looked at assume, the former explicitly and the latter implicitly, that any action is or can be learned (It is, furthermore rather difficult to avoid this assumption, since the categories "learned" and "innate" appear to be exhaustive). It is only one step further to suggest that this learning can take place within the educational system, and the statements admit of no limits to the capacity of educational systems. It follows from these considerations, however, that if one is to treat desired actions as goals, a general classification of action of the Parsonian type would be relevant to the analysis of goal statements.

It should, however, be pointed out that those preference statements about outcomes which we can identify as referring to action or action patterns are in a minority. Residually we can identify at least two more types of statement.

A) Those which do not mean anything (i.e. those statements which have no operational specification or can be given none)

This category can be subdivided into:-

i) general statements which judge the educational system to be inadequate, but are not sufficiently explicit for us to be able to identify any of the elements which are inadequate.

e.g. "Better high school education is essential both to raise the calibre of students entering college and to provide the best training possible for those who end education with the twelfth grade".

(N.B. Statements such as this should be treated with extreme caution even where the terms can be given operational specification, since the internal logic of the statements implies the inadequacy of goal achievement, and hence presupposes the outcome of the whole project which is being undertaken).

ii) statements involving values (Which, sociological theory to the contrary, do not appear to be operational terms) These are most frequently found as relating education to "democracy", "liberty", "freedom", etc. We simply do not know what is meant by statements such as "To develop for the regulation of one's personal civic life a code of behavior based on ethical princeples consistent with democratic ideals."

B) Those which deal only with a psychological variable (characteristic of an actor)

This category can again be subdivided

i) those where the characteristics stated coincide with dimensions of personality and thus can be conceptualized in terms of existing psychology

e.g. statements about "motivation"
statements about "intellectual capacities"

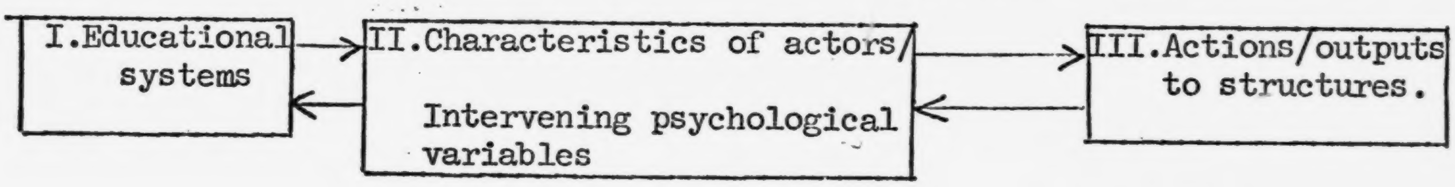
ii) those where there is no such coincidence

e.g. "developing personality"

(The latter category again may merely refer to those statements about the psychological make up of actors to which we can give no more explicit specification, i.e. to those statements which do not mean anything.)

---

3. It seems that we are dealing with the following general schema.


I.Educational systems
II.Characteristics of actors/ Intervening psychological variables
III.Actions/outputs to structures.

The logic of most preference statements proceeds from right to left, but the actual statements may be made to refer to any of the boxes. (In general, a statement which is logically related to another to the

the right will be found in the original classification as a means statement) The educational process proceeds from left to right.

The problem now is quite obviously "How do we proceed?" The alternatives appear to be to develop a classification scheme based on either box I. or box II. The former looks remarkably like "personality theory", the latter remarkably like "general action theory", both are fairly obviously problematic . . . .(As one test of whether we are dealing with a classification scheme involving "any action" we tried to think of an action which we could not impute to educational systems as a goal statement of the type with which we have been dealing. The only actions which we could not view as goals appeared to be in some sense "unlearned" and to fall outside the explanatory framework of sociology as a science.)

AN EXPLANATION OF THE FILE

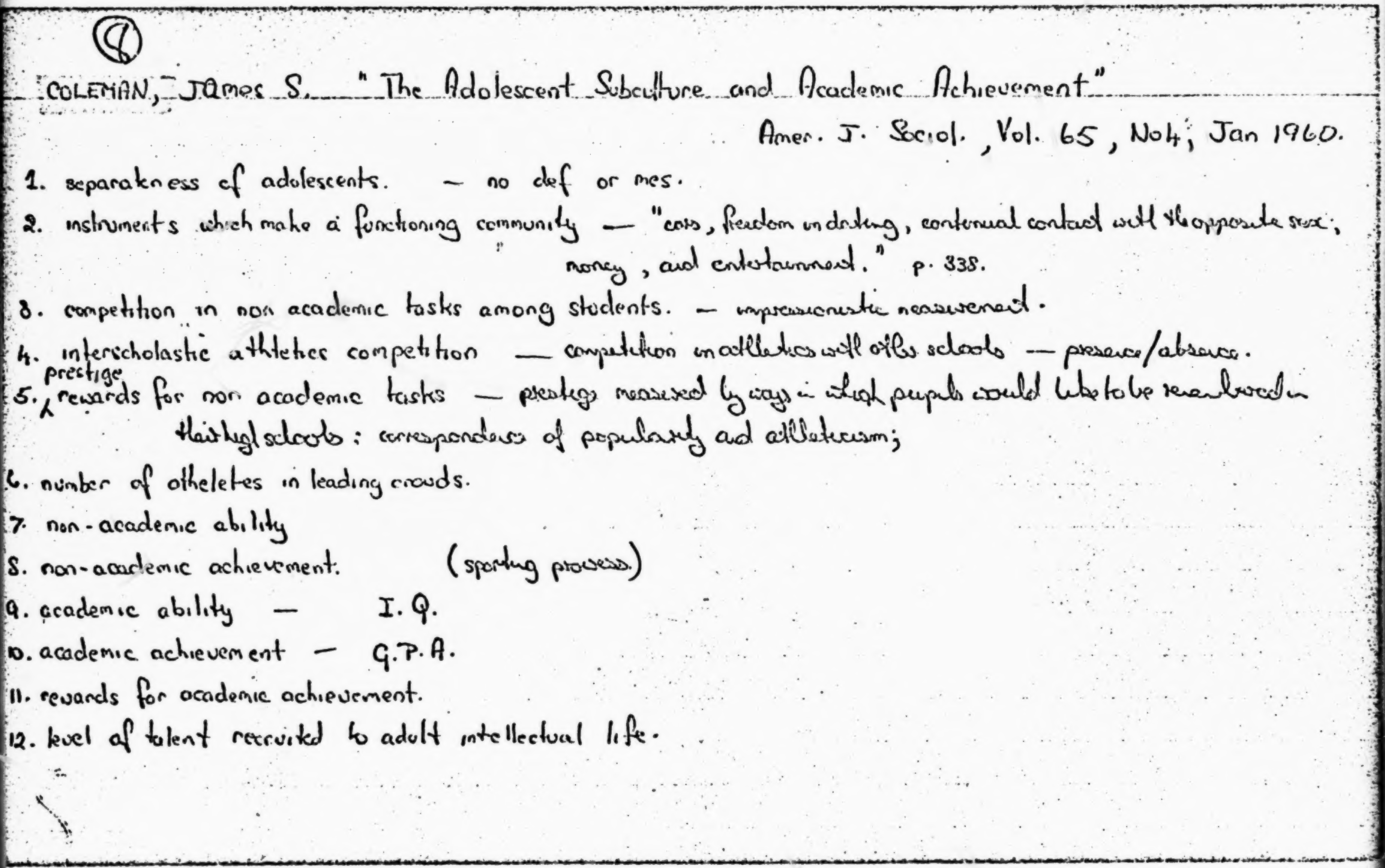
The file contains three sorts of colored cards, red, blue and white. The red and blue cards contain the summaries of the articles and books, and a green marking along the top edge of any card indicates a book source. The white cards are individual variable cards.

Summaries of Articles

Since the articles themselves tend to be more systematically expressed than the books, those cards summarizing articles include most information. Each article has at least two cards, one red, one blue: some long or excessively complex articles have two or three cards of each color. On the front of the red card appears a list of the variables, together with definitions of these variables (where given or implied) and measures used (where measurement was made). On the back of the red card is a brief description of the sample and methodology used in the article. On the front of the blue card appears the theory and relationships between variables considered in the article. These relationships are expressed symbolically and explained in the text above. The numbering of the variables corresponds to the numbering on the pink card. On the back of the blue card is a summary of the evidence for the relationships stated schematically on the front of this card.

The following is an example of an article summarized.

RED CARD - FRONT



⑨

COLEMAN, James S. "The Adolescent Subculture and Academic Achievement"

Amer. J. Sociol., Vol. 65, No 4, Jan 1960.

1. separateness of adolescents. — no def or mes.
2. instruments which make a functioning community — "cars, freedom in dating, continual contact with the opposite sex, money, and entertainment." p. 338.
3. competition in non academic tasks among students. — impressionistic measurement.
4. interscholastic athletic competition — competition in athletics with other schools — presence/absence.
5. prestige rewards for non academic tasks — prestige measured by ways in which pupils would like to be remembered in their high schools: correspondence of popularity and athleticism;
6. number of athletes in leading crowds.
7. non-academic ability
8. non-academic achievement. (sporting prowess)
9. academic ability — I.Q.
10. academic achievement — G.P.A.
11. rewards for academic achievement.
12. level of talent recruited to adult intellectual life.

RED CARD - BACK

METHODOLOGY -

Sample — pupils from 10 high schools in the Midwest, all coeducation except 1 Catholic boys school.

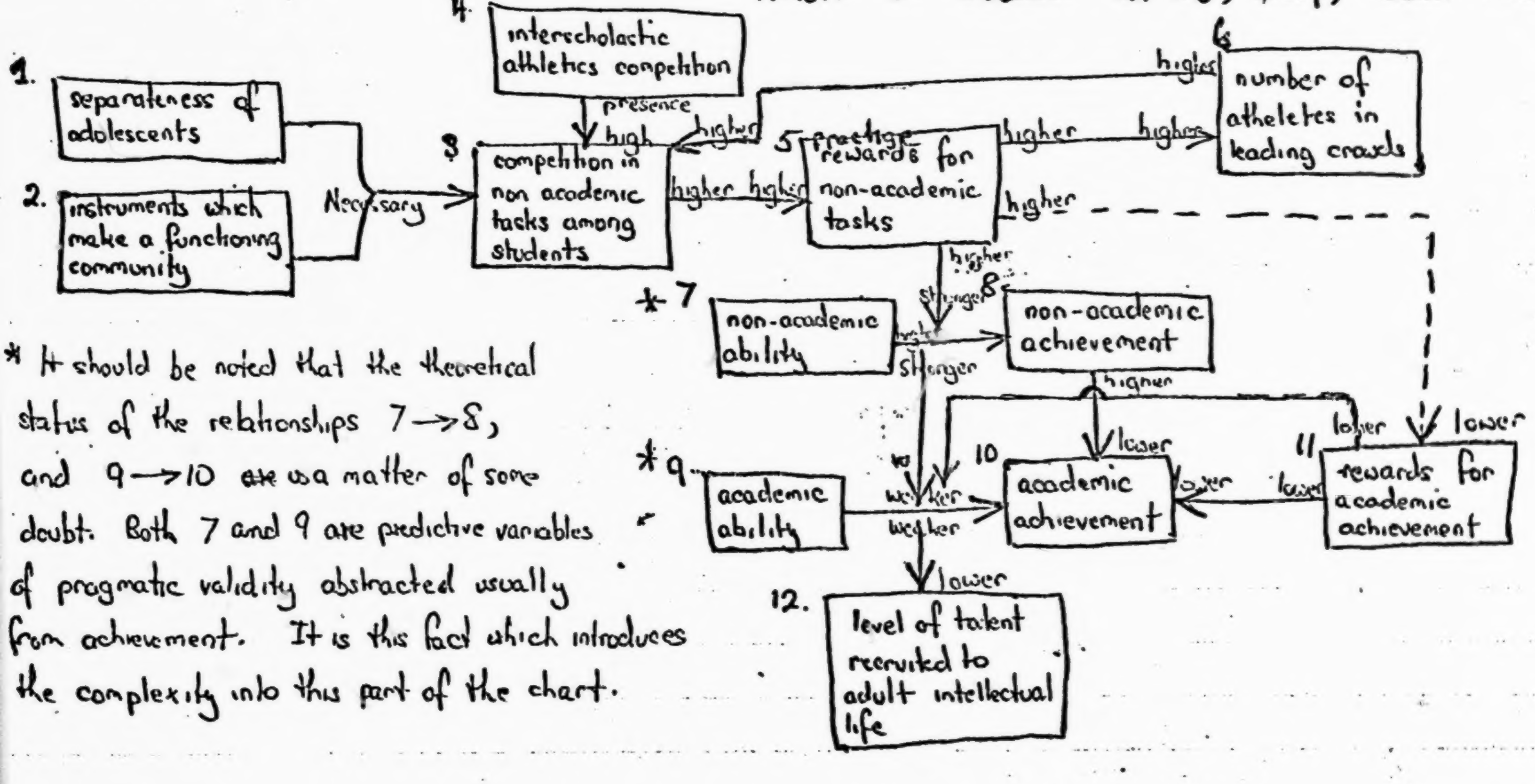
Method. Boys asked to name best athletes, best student & boy most popular with girls. and to state who were "members of the leading crowd." Girls similarly questioned with "leader in extracurricular activity" substituted for athletic star. All asked "how they would like to be remembered."

Aca. Data on academic achievement (grades) and IQ. also taken from school records.

BLUE CARD - FRONT

④

COLEMAN, James S. "The Adolescent Subculture and Academic Achievement"
Amer. J. Sociol. Vol 65, No.4, Jan. 1960.


4.
interscholactic athletics competition
presence
high
1.
separateness of adolescents
2.
instruments which make a functioning community
Necessary
3
competition in non academic tasks among students
higher
higher higher
5
~~prestige~~
rewards for non-academic tasks
higher
higher
6
higher
number of atheletes in leading crowds
higher
higher
strong
*7
non-academic ability
weaker
stronger
8
non-academic achievement
higher
lower
lower
lower
11
rewards for academic achievement
lower
lower
10
academic achievement
*9.
academic ability
weaker
weaker
lower
12.
level of talent recruited to adult intellectual life

* It should be noted that the theoretical status of the relationships 7 → 8, and 9 → 10 are is a matter of some doubt. Both 7 and 9 are predictive variables of pragmatic validity abstracted usually from achievement. It is this fact which introduces the complexity into this part of the chart.

BLUE CARD - BACK

1, 2 → 3. impressionistic evidence from American society.

4 → 3. established by comparing those schools which competed in athletics with two that did not. Two that did not showed higher perform emphasis on academic tasks.

3 → 5 assumption. + impressionistic evidence.

5 → 6
6 → 3 } number of athletes measured, but other 5 & 3 measured by intuitive assessment only. The suspicion that independent measurement does not take place here is probably justified, and if evidence for 5 & 3 is drawn from 6 the evidential basis for this subset is tautological.

5 → 7, 8 This subset involves the assertion that where rewards for academic performance are high, a strong relationship will exist between athletic ability & performance. Since athletic ability is not measured no evidence is actually given.

5 → 11. This appears to be implicit in much of the argument. No evidence is presented and the logical basis may be dubious.

7 ↓ 8 → 9 ↓ 10 The stronger the relation between non-academic ability and performance the weaker the relation between academic ability & performance. The first relationship is assumed to be strong in the schools under study. For the second students with high G.P.A. are taken and their IQ's compared with the IQ distribution of their schools to determine whether high achievers are the brightest students in the schools.

11 → 9 ↓ 10 The above is then backed up with evidence which by comparing schools shows that the more the emphasis on academic performance in the subculture, the stronger the relationship between 9 & 10.

11 → 10 no evidence — problem is comparability of G.P.A.s between schools ∴ two above are used

10 — 12 asserted.

Individual Variable Cards

Once the task of summarizing was completed each variable was copied onto a separate white card (see below) and repeated for as many entries as were thought necessary in an alphabetized file. Examples of two of the variables from the article by Coleman above are given below. The punched holes along the top of this card indicate classification of the variable into "input", "system characteristic", or "outcome".

WHITE CARD - VARIABLE (from Coleman)

Academic ability

V: academic ability - ~~IQ~~. Intelligence Quotient
S: pupils from 10 high schools in midwest.

(8) COLEMAN, James S.

WHITE CARD - VARIABLE (from Coleman)

Separateness

V: separateness of adolescents
S: (pupils from 10 high schools in midwest) adolescents

(8) COLEMAN, James S.

Summaries of Books

We began this summary using the same format as we had used for the articles. It was, however, quickly discovered that since most of the books are speculative and non-empirical most of the format was redundant. We could not add "definition" or "measurement" to our list of variables, nor describe the methodology, nor summarize evidence for relationships since none of these existed. Three of the four sides of our two cards thus came to be unused, and with the books we came to take down only the propositions in their symbolic form, noting evidence beneath the propositions on the rare occasions on which evidence was given. Book summaries thus include for the most part only schematically represented propositions and the color of the card (Both red and blue are used) is irrelevant.

REFERENCES: ARTICLES

I Articles

Bonjean, Charles M., Reece, McGee.

"Scholastic Dishonesty Among Undergraduates in Different Forms of Social Control"

Sociol. of Educ., Vol. 38, No. 2, Winter 1965.

Boyle, R.P.

"Effect of High School on Student Aspirations"

Amer. J. Sociol., Vol. 71, No. 6, May 1966.

Brookover, Wilbur B., Shailer, Thomas, Paterson, Ann.

"Self Concept of Ability and School Achievement"

Sociol. of Educ., Vol 37, No. 3, Spring 1964.

Caplan, Stanley W., Ruble, Ronald A.

"A Study of Culturally Imposed Factors on School Acheivement in a Metropolitan Area"

J. Educ. Research, Vol. 58, No. 1, Sept.1964.

Caro, Francis G.

"Deferred Gratification, Time Conflict, and College Attendance"

Sociol. of Educ., Vol. 38, No. 4, Summer 1965.

Case, Herman M.

"College as a Factor in Occupational Choice: A Study of Different Perceptions by Farm and Non-Farm Youth"

J. Educ. Sociol., Vol. 30, No. 4, Dec. 1964.

Cervantes, Lucius F.

"Family Background, Primary Relationships, and the High School Dropout"

J. of Marriage and the Family, Vol. 27, No. 2, May 1965.

Clark, Burton R.

"The Cooling-out Function in Higher Education"

Amer. J. Sociol., Vol. 65, No. 6, May 1960.

Clark, Kenneth B.

"Segregated Schools in New York City"

J. Educ. Sociol., Vol. 36, No. 6, Feb. 1963.

Cline, Victor B., Richards, James, Needham, Walter E.,

"Factor Analysis of Self Ratings, Teacher Ratings, and Indices of Achievement in High School Science"

J. Educ. Res., Vol. 58, No. 1, Sept. 1964.

Cohen, Elizabeth G.

"Parental Factors in Educational Mobility"

Sociol. of Educ., Vol. 38, No. 5, Fall 1965.

Coleman, James S.

"The Adolescent Subculture and Academic Achievement"

Amer. J. Sociol., Vol. 65, No. 4, Jan. 1960.

Dreeben, Robert

"The Contribution of Schooling to the Learning of Norms"

Harvard Educ. Rev., Vol. 37, No. 2, Spring 1967.

Duncan, Beverly

"Education and Social Background"

Amer. J. Sociol., Vol. 72, No. 4, Jan. 1967.

Eckland, Bruce K.

"Social Class and College Graduation: Some Misconceptions Corrected"

Amer. J. Sociol., Vol. 70, No. 1, July, 1964.

Eckland, Bruce K.

"Academic Ability, Higher Education, and Occupational Mobility"

Amer. Sociol. Rev., Vol. 30, No. 5, Oct. 196[illegible].

Elder, Glen H., Jr.

"Achievement Orientations and Career Patterns of Rural Youth"

Sociol. of Educ., Vol. 37, No. 1, Fall 1963.

Elder, Glen H.

"Life Opportunity and Personality: Some Consequences of Stratified Secondary Education in Great Britain"

Sociol. of Educ., Vol. 38, No. 3, Spring, 1965.

Elley, Warwick B.

"Attitude Change and Education for International Understanding"

Sociol. of Educ., Vol. 37, No. 4, Summer 1964.

Glick, Paul C., Miller, H.P.

"Educational Level and Potential Income"

Amer. Sociol. Rev., Vol. 21, No. 3, June 1956.

Glick, Paul C., Carter, H.

"Marriage Patterns and Educational Level"

Amer. Sociol. Rev., Vol. 23, No. 3, June 1958.

Goren, Arnold L.

"Student Activities and the Teacher"

J. Educ. Sociol., Vol. 29, No. 8, April 1956.

Hansen, Carl F.

"Scholastic Performance of Negro and White Pupils in Washington, D.C."

J. Educ. Sociol., Vol. 36, No. 6, Feb. 1963.

Herriott, Robert E.

"Some Social Determinants of Educational Aspiration"

Harvard Educ. Rev., Vol. 33, No. 2, Spring 1963.

Litt, Edgar

"Civic Education, Community Norms, and Political Indoctrination"

Amer. Sociol. Rev., Vol. 28, No. 1, Feb. 1963.

Lynn, R.

"The Relation between Educational Achievement and School Size"

British J. Sociol., Vol. 10, No. 2, June 1959.

McDill, Edward L., Coleman, James S.

"High School Social Status, College Plans, and Interest in Academic Achievement: A Panel Analysis"

Amer. Sociol. Rev., Vol. 28, No. 6.

Miller, Jerry.L.L.

"Occupational Choice and the Educational System"

J. Educ. Sociol., Vol. 34, No. 3, Nov. 1960.

Sharp, Laure

"The Meaning of the Bachelor's Degree"

Sociol. of Educ., Vol. 37, No. 2, Winter 1963.

Schroder, Ralph

"Academic Achievement of the Male College Student"

J. of Marriage and Family Living, Vol. 25, No. 4, Nov. 1963.

Schultz, Raymond E., Merle, H. Olsen

"Interest Patterns of Best and Poorest Teachers"

J. Educ. Sociol., Vol. 29, No. 3, Nov. 1955.

St. John, Nancy Hoyt

"Effect of Segregation on Aspirations of Negro Youth"

Harvard Educ. Rev., Vol. 36, No. 3, Summer 1966.

Timascheff, N.S.

"Higher Education for What?"

J. Educ. Sociol., Vol. 29, No. 2, Oct. 1955.

Tumin, Melvin M., Barton, P., Burrus, B.

"Education, Prejedice and Desegregation"

Amer. Sociol. Rev., Vol. 23, No. 1, Feb. 1958.

Wallace, Walter L.

"Institutional and Life Cycle Socialization of College Freshmen"

Amer. J. Sociol., Vol. 70, No. 3, Nov. 1964.

Weinberg, Carl

"Family Background and Deviance or Conformity to School Expectations"

J. of Marriage and Family, Vol. 26, No. 1, Feb. 1964.

Westoff, Charles F., Potvin, Raymond

"Higher Education, Religion and Women's Family Size Orientations"

Amer. Sociol. Rev., Vol. 31, No. 4, Aug. 1966.

Wright, Charles R.

"Success or Failure in Earning Graduate Degrees"

Sociol. of Educ., Vol. 38, No. 1, Fall 1964.

2. Books

Anderson, Arnold, C., and Bowman, Mary Jean.

'Education and Economic Development

Chicago 1965, Aldine Publ. Co.

Bestor, Arthur

The Restoration of Learning

New York, 1956, Alfred A. Knopf.

Clark, Burton R.

Educating the Expert Society

San Francisco, 1962, Chandler Publ. Co.

Coleman, James S.

Education and Political Development

Princeton 1965, by Princeton University Press

Conant, J.B.

Slums and Suburbs

New York, by McGraw Hill Book Co.

Conant, James Bryant

The Child, the Parent and the State

Cambridge, 1960, by Harvard University Press

Conant, James Bryant

Education in a Divided World

Cambridge, 1949 by Harvard University Press

Cramer, John F., Brown, George S.

Contemporary Education: A Comparative Study of National Systems

New York, 1956, Harcourt Brace and Co.

Dahlke, H. Otto

Values in Culture and Classroom

New York, 1958, Harper and Brothers

Fichter, Joseph

The Parochial School

Notre Dame, 1958, University of Notre Dame Press

Friedenberg, Edgar Z.

The Vanishing Adolescent

Boston, 1964, by Beacon Press.

Goodman, Paul

Compulsory Mis-education

New York, 1964, Horizon Press.

Harbison, Frederick, Myers, Charles A.

Education Manpower and Economic Growth

New York, 1964, McGraw Hill.

Kallenbach, W. Warren, Hodges, Harold M.

Education and Society

Columbus, Ohio, 1963, Charles E. Merrill.

Mannheim, Karl, Stewart, C.A.

An Introduction to the Sociology of Education

1962, Routledge, Kegan, Paul.

Musgrave, P.W.

The Sociology of Education

London, 1966, by Methuen

National Society for the Study of Education

Individualizing Instruction

Chicago, 1962, NSSE.

National Educational Association

Education in a Changing Society

Washington, D.C., 1963, NEA

Neill, A.S.

Summerhill, A Radical Approach to Child Rearing

N.Y., 1960, Hart Publ. Co.

Ottoway, A.K.C.

Education and Society

London, 1962, Routledge, Kegan, Paul

Patterson, Franklin K. et al.

The Adolescent Citizen

N.Y., Free Press.

Sanford, Nevitt

The American College

N.Y., 1962, Wiley

Stember, C.H.

The Effect of Education on Prejudice against Minority Groups

N.Y., 1960, American Jewish Comm.

xxx

EXAMPLES OF CLASSIFIED VARIABLES

KNOWING

ECONOMIC

12	A) REQUIREMENTS OF ROLE	ability to earn living	2
46	A) REQUIREMENTS OF ROLE (ii) knowledge	success/failure in college curriculum germane to occupational choice	3
383	A) REQUIREMENTS OF ROLE (iii) skills	industrial skills	3
588	A) REQUIREMENTS OF ROLE (iii) skills (a) mental	number of adequate persons available for research	2

POLITICAL

113	A) REQUIREMENTS OF ROLE (ii) knowledge	political awareness	7

PRIMARY GROUP

689	A) REQUIREMENTS OF ROLE (iii) skills (a) mental	inter-sex understanding	2
362	B) STANDARDS AND NORMS	recognition of home as source of values and standards	1

CULTURAL HERITAGE

345		perception of increased probability of attending graduate school	1
17, 34, 108, 391	A) REQUIREMENTS OF ROLE (ii) knowledge	17 - academic achievement (grade point average 34 - level of education achieved 108 - cultural awareness 391 - lack of presence of information	62

KNOWING

CULTURAL HERITAGE

2, 9, 373, 468	A) REQUIREMENTS OF ROLE (iii) skills (a) mental	2 - ability to analyze 9 - ability to generalize 373 - imagination 468 - linguistic skills	21
24	A) REQUIREMENTS OF ROLE (iii) skills (b) physical	non-academic achievement (sporting prowess)	1
185	B) STANDARDS AND NORMS	common sense-science criteria for accepting evidence	1

ORIENTATIONS TO SELF AND OTHERS

15, 111		15 - self-concept of ability 111 - awareness of persons other than self and immediate primary groups	4
15.1	A) REQUIREMENTS OF ROLE (i) relational (a) interpersonal	capability for social relations	6
114, 122-145, 563	A) REQUIREMENTS OF ROLE (ii) knowledge	114 - social awareness 122-145 - racial prejudice 563 - likelihood of opinion change by students	33
530	A) REQUIREMENTS OF ROLE (iii) skills	new skills	1
11, 186, 203	A) REQUIREMENTS OF ROLE (iii) skills (a) mental	11 - leadership ability 186 - essential communication skills 203 - creativity	6
676	A) REQUIREMENTS OF ROLE (iv) resources (c) psychic	development of self definition	3

KNOWING

ORIENTATIONS TO SELF AND OTHERS			
459	B) STANDARDS AND NORMS	learning of norms	3
775	B) STANDARDS AND NORMS (i) moral and legal	understanding of values	3
115	C) SPECIFICATIONS OF ROLE (ii) time allocations	awareness of times other than immediate presence	1
112	C) SPECIFICATIONS OF ROLE (iii) space and movement	awareness of places other than own locality	2
65	D) REWARDS OF ROLE (iii) psychic (b) interpersonal	minority-group sense of personal humiliation	4

VALUING

ECONOMIC

93	A) REQUIREMENTS OF ROLE (iv) resources (c) psychic	occupational aspirations	4
28	B) STANDARDS AND NORMS (i) moral and legal	acceptance of work discipline	1
507	D) REWARDS OF ROLE (i) property	likelihood of valuing job for money rather than "prospects"	1

POLITICAL

405	A) REQUIREMENTS OF ROLE (i) relational (b) intergroup	interest in national affairs	4
102	A) REQUIREMENTS OF ROLE (iii) skills (a) mental	abstract attitude in political sphere	1
171	A) REQUIREMENTS OF ROLE (iv) resources (c) psychic	development of citizenship responsibility	2
101, 105, 233, 571	B) STANDARDS AND NORMS (i) moral	101 - political attitudes 105 - minority group disrespect of authority 233 - democratic values 571 - punishment-reform in penological theory	14

PRIMARY GROUP

308	A) REQUIREMENTS OF ROLE (iv) resources (b) personnel	desired family size	1

SOCIALIZATION

83	B) STANDARDS AND NORMS (i) moral and legal	approval of school integration	2

VALUING

SOCIALIZATION

238	D) REWARDS OF ROLE (iii) psychic	felt disappointment of prospect (of not completing college)	2

CULTURAL HERITAGE

22		proportion of students who value scholastic attainment	1
18	A) REQUIREMENTS OF ROLE (ii) knowledge	valuation on scholastic achievement, positive/negative	6
92, 192	A) REQUIREMENTS OF ROLE (iv) resources (c) psychic	92 - educational aspiration 192 - serious concern for ideas	15
590	B) STANDARDS AND NORMS (i) moral and legal	traditionalism-secularism in social philosophy	2

ORIENTATIONS TO SELF AND OTHERS

197		non-conservatism	1
754	A) REQUIREMENTS OF ROLE (i) relational	tolerance	1
104, 472	A) REQUIREMENTS OF ROLE (i) relational (a) interpersonal	104 - tolerant attitude toward persons 472 - loyalty	5
404, 407	A) REQUIREMENTS OF ROLE (i) relational (b) intergroup	404 - interest in minority group problems 407 - international understanding	4
58	A) REQUIREMENTS OF ROLE (iii) skills (a) mental	achievement motivation	2

VALUING

ORIENTATIONS TO SELF AND OTHERS			
74, 80	A) REQUIREMENTS OF ROLE (iv) resources (c) psychic	74 - collective lack of motivation 80 - anxiety	33
27	B) STANDARDS AND NORMS	accepting of norms	5
59, 99	B) STANDARDS AND NORMS (i) moral and legal	59 - achievement norms 99 - liberal attitude on social issues	14
240	C) SPECIFICATIONS OF ROLE (i) routines	acceptance of work discipline	1
26	C) SPECIFICATIONS OF ROLE (ii) time allocations	acceptance of change	3
76	D) REWARDS	socio-economic ambition	1
659	D) REWARDS (ii) psychic (a) intrapersonal	satisfaction	1

DOING

ECONOMIC

61, 161		61 - occupational achievement 161 - choice of occupation	3
389	A) REQUIREMENTS OF ROLE	occupational inflexibility	1
540	A) REQUIREMENTS OF ROLE (i) relational	job adjustment	1
198	A) REQUIREMENTS OF ROLE (iii) skills (a) mental	discrimination among consumer products	2
486	A) REQUIREMENTS OF ROLE (iv) resources (b) human	shortage of scientific manpower	3
179, 348	C) SPECIFICATIONS OF ROLE (i) routines	179 - clerical work 348 - employment full time at graduation	5
543	D) REWARDS OF ROLE	selection of occupation involving personal gain	2
545	D) REWARDS OF ROLE (i) property	low income or not	2
562	D) REWARDS OF ROLE (iii) psychic (b) interpersonal	occupational status	1

POLITICAL

187	A) REQUIREMENTS OF ROLE (i) relational (b) intergroup	involvement of students in struggle against world communism	1
800	B) STANDARDS AND NORMS (i) moral and legal	likelihood of voting	3

DOING

POLITICAL

267, 170	C) SPECIFICATIONS OF ROLE (i) routines	167 - civic participation and community projects	4
188	D) REWARDS OF ROLE (iii) psychic gratifications	success in the fight against world communism	1

PRIMARY GROUP

160, 483	A) REQUIREMENTS OF ROLE (i) relational (a) interpersonal	160 - rigidity-permissiveness in child-rearing patterns 483 - extent of marital stability	7
226	A) REQUIREMENTS OF ROLE (iii) skills (a) mental	formality/deliberation of marriage planning or not	1
309	A) REQUIREMENTS OF ROLE (iv) resources (a) material	requirement of support from man by family-of-procreation	1
476	A) REQUIREMENTS OF ROLE (iv) resources (b) human	number of marriages by a person	1
806	C) SPECIFICATIONS OF ROLE (i) routines	widowhood-other	5
69	C) SPECIFICATIONS OF ROLE (ii) time allocations	median age at divorce	11

SOCIALIZATION

747	A) REQUIREMENTS OF ROLE (iv) resources (b) human	number of talented persons available for teaching	1

CULTURAL HERITAGE

377	A) REQUIREMENTS OF ROLE (i) relational (a) interpersonal	recreational development of individual	1

DOING

CULTURAL HERITAGE			
394	A) REQUIREMENTS OF ROLE (ii) knowledge	application of knowledge to new contexts	1
204	C) SPECIFICATIONS OF ROLE (i) routines	creativity and transmission of research techniques	1
ORIENTATIONS TO SELF AND OTHERS			
67, 106		67 - adult life 106 - autonomy and freedom	4
252, 473	A) REQUIREMENTS OF ROLE (i) relational	252 - social dominance 473 - lying	3
66, 380	A) REQUIREMENTS OF ROLE (i) relational (a) interpersonal	66 - social adjustment 380 - individuality	12
87	A) REQUIREMENTS OF ROLE (iii) skills (a) mental	utilization of potential	2
177	A) REQUIREMENTS OF ROLE (iv) resources (c) psychic	minority-group members of middle class withdrawel and submissive behavior	4
629	B) STANDARDS AND NORMS	rebellious behavior	2
227	B) STANDARDS AND NORMS (i) legal moral	rate of juvenile delinquency	1
239	C) SPECIFICATIONS OF ROLE (i) routines	practice in self-discipline	1
PHYSICAL WELL-BEING			
358, 591		358 - physical health 591 - vigor	9

THE OUTCOME CLASSIFICATION SCHEME

	DOMAINS						
NOWING	Economic	Political	Primary Group	Socialization and/or Education	Activities Involving Cultural Heritage	Orientations to Self and Others	Physi cal Well- bein
Requirements of Role	2						
1. Relational							
a. Inter-Personal						6	
b. Inter-Group							
2. Knowledge	3	7			62	33	
3. Skills	3					1	
a. Mental	2		2		21	6	
b. Physical					1		
4. Resources							
a. Material							
b. Human							
c. Psychic						3	
Standards and Norms			1		1	3	
1. Moral and Legal						3	
2. Aesthetic							
3. Style and Manners							
Specifications of Role							
1. Routines							
2. Time-allocations						1	
3. Space and Movement						2	
Rewards of Role							
1. Property							
2. Power							
3. Psychic							
a. Intra-Personal							
b. Inter-Personal						4	
(Unclassified)					1	4	

* This Domain is not

VALUING	Economic	Political	Primary Group	Socialization and/or Education	Activities Involving Cultural Heritage	Orientations to Self and Others	Phy cal Well bein
			DOMAINS				
Requirements of Role							
1. Relational						1	
a. Inter-Personal						5	
b. Inter-Group		4				4	
2. Knowledge					6		
3. Skills							
a. Mental		1				2	
b. Physical							
4. Resources							
a. Material							
b. Human			1				
c. Psychic	4	2			15	33	
Standards and Norms						5	
1. Moral and Legal	1	14		2	2	14	
2. Aesthetic							
3. Style and Manners							
Specifications of Role							
1. Routines						1	
2 Time-allocations						3	
3. Space and Movement							
Rewards of Role						1	
1. Property	1						
2. Power							
3. Psychic				2			
a. Intra-Personal						1	
b. Inter-Personal							
(Unclassified)					1	1	

DOING	Economic	Political	Primary Group	Socialization and/or Education	Activities Involving Cultural Heritage	Orientations to Self and Others	Ph ca We be
	DOMAINS						
Requirements of Role	1						
1. Relational	1					3	
a. Inter-Personal			7		1	12	
b. Inter-Group		1					
2. Knowledge					1		
3. Skills							
a. Mental	2		1			2	
b. Physical							
4. Resources							
a. Material			1				
b. Human	3		1	1			
c. Psychic						4	
Standards and Norms						2	
1. Moral and Legal		3				1	
2. Aesthetic							
3. Style and Manners							
C. Specifications of Role							
1. Routines	5	4	5	1		1	
2. Time-Allocations			11				
3. Space and Movement							
D. Rewards of Role	2						
1. Property	2						
2. Power							
3. Psychic		1					
a. [illegible]-Personal							
b. Inter-Personal	1						
(Unclassified)	3					4	

A SELECTED BIBLIOGRAPHY OF ARTICLES AND BOOKS DIGESTED FOR THE PURPOSES OF THE STUDY

Goals of Education:

Beezley, P. C., *Education for What?* New York: Bookmailer, 1963.

Bestor, Arthur, *The Restoration of Learning*. New York: Alfred A. Knopf, 1956.

Bloom, Benjamin S., et al, *Taxonomy of Educational Objectives, I. The Cognitive Domain.* New York: David McKay Company, Inc., 1956.

Booth, Wayne C. ed., *The Knowledge Most Worth Having.* Chicago: The University of Chicago Press, 1966.

Brim, Orville C., Jr., *Sociology and the Field of Education.* New York: Russell Sage Foundation, 1958.

Bruner, Jerome S., *The Process of Education.* Cambridge, Mass: Harvard University Press, 1966.

Bruner, Jerome S., *Toward a Theory of Instruction* Cambridge, Mass: Harvard University Press, 1966.

Butler, Nicholas Murray, *The Meaning of Education: Contributions to a Philosophy of Education*. New York: Charles Scribner's Sons, 1915.

Clark, Burton R., *Educating the Expert Society.* San Francisco: Chandler Publishing Co., 1962.

Commission on the Reorganization of Secondary Education, "Cardinal Principles of Secondary Education." Bureau of Education *Bulletin*, No. 35. 1918

Conant, James B., *The American High School Today.* New York: McGraw-Hill Book Company, Inc., 1959.

Conant, James B., *Shaping Educational Policy.* New York: McGraw-Hill Book Company, 1964.

Cremin, Lawrence A. *The Transformation of the School: Progressivism in American Education, 1876-1967.* New York: Alfred A. Knopf, 1961.

Cubberley, Ellwood P., *Public Education in the United States: A Study and Interpretation of American Educational History.* Boston: Houghton Mifflin Company, 1919.

Dewey, John, "My Pedagogic Creed", E. L. Kellogg, 1897. In Archambault, Reginald D., John Dewey on Education: Selected Writings. New York: The Modern Library, 1964.

Educational Policies Commission, The Central Purpose of American Education. Washington, D. C.: National Education Association, 1961.

Educational Policies Commission, The Contemporary Challenge to American Education. Washington, D. C.: National Education Association, 1958.

Educational Policies Commission, Contemporary Issues in Elementary Education. Washington, D. C.: National Education Association, 1960.

Educational Policies Commission, Education and the Disadvantaged American. Washington, D. C.: National Education Association, 1962.

Educational Policies Commission, An Essay on Quality in Public Education. Washington, D. C.: National Education Association, 1959.

Educational Policies Commission, The Purposes of Education in American Democracy. Washington, D. C.: National Education Association, 1938.

Educational Policies Commission, Universal Opportunity for Education Beyond The High School. Washington, D. C.: National Education Association, 1964.

Fine, Benjamin, Our Children are Cheated. New York: Henry Holt and Company, 1947.

French, Will et al., Behavioral Goals of General Education in High School. New York: Russell Sage Foundation, 1957.

Gardner, John W., Excellence. New York: Harper & Row, Publishers, 1961.

Gardner, John W., Self-Renewal. New York: Harper & Row, Publishers, 1964.

Ginzberg, Eli., ed., The Nation's Children: 2: Development and Education. New York: Columbia University Press, 1960.

Goodlad, John I., ed., The Changing American School. The Sixty-fifth Yearbook of the National Society for the Study of Education, Part II. Chicago: National Society for the Study of Education, 1966.

Goodman, Paul, Contemporary Mis-Education. New York: Horizon Press, 1964.

Hambridge, Gove, New Aims in Education. New York: McGraw-Hill Book Company, Inc., 1940.

Hanna, Paul R., ed., Education: An Instrument of National Goals. New York: McGraw-Hill Book Company, 1962.

Harvard University Committee on the Objectives of a General Education in Free Society, General Education in a Free Society. Cambridge, Mass: Harvard University Press, 1945.

Hechinger, Fred M., Pre-School Education Today. Garden City, New York: Doubleday & Company, Inc., 1966.

Hook, Sidney, "Perennial and Temporal Goals in Education". The Journal of Higher Education, Vol. XXIII, No. 1, January, 1952, pp 1 - 12.

Hutchins, Robert Maynard, Education for Freedom. Baton Rouge, La.: Louisiana State University Press, 1946.

Kearney, Nolan C., Elementary School Objectives. New York: Russell Sage Foundation, 1953.

Koerner, James D., The Case for Basic Education. Boston: Little, Brown and Company, 1959.

Krathwohl, David R., et al, Taxonomy of Educational Objectives. II. The Affective Domain. New York: David McKay Comapny, Inc., 1964

Lindvall, C. M., Defining Educational Objectives. Pittsburgh: University of Pittsburgh Press, 1964.

Maritain, Jaques, Education at the Crossroads. New Haven: Yale University Press, 1943.

Members of the Faculties of Andover, Exeter, Lawrenceville, Harvard, Princeton, and Yale, General Education in School and College. Cambridge, Mass: Harvard University Press, 1952.

Minnesota University Graduate School, Social Science Research Center, Perspectives on the Conant Report. Minneapolis, Minn: University of Minnesota, 1960.

National Education Association Project on Instruction, Schools for the Sixties. New York: McGraw-Hill Book Company, 1963.

President's Commission on National Goals, Goals for Americans. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1960.

Rickover, H. G., Education for All Children. Washington, D. C.: U. S. Government Printing Office, 1962.

Rockefeller Brothers Fund, The Pursuit of Excellence: Education and the Future of America. Special Studies Project V of America at Mid-Century Series. Garden City, New York: Doubleday & Company, 1958.

Thwing, Charles Francis, What Education Has the Most Worth? New York: The Macmillan Company, 1924.

United States Office of Education, Contemporary Issues in American Education. Washington, D. C.: U. S. Government Printing Office, 1965; or, U. S. Office of Education Bulletin, No. 3, 1966.

U. S. President's Commission on Higher Education, Higher Education for American Democracy, I, Establishing the Goals. Washington, D. C.: U. S. Government Printing Office, 1947.

Whitehead, Alfred North, The Aims of Education. New York: The Macmillan Company, 1929.

Evaluation

Ahmann, J. Stanley, and Glock, Marvin D., Evaluating Pupil Growth. Boston: Allyn and Bacon, Inc., 1963.

Aikin, Wilford M., The Story of the Eight Year Study. New York: Harper & Brothers, 1942.

American Council on Education Commission on Accreditation of Service Experiences, A Guide to the Evaluation of Educational Experiences in the Armed Services. Washington, D. C.: American Council on Education, 1954.

American Council on Education, The Executive Committee of the Cooperative Study in General Education, Cooperation in General Education. Washington, D. C.: American Council on Education, 1947.

Barton, Allen H., Organizational Measurement and Its Bearing on the Study of College Environments. New York: College Entrance Examination Board, 1961.

Barton, Allen H., Studying the Effects of College Education: A Methodological Evaluation of Changing Values in College. New Haven, Conn.: The Edward W. Hazen Foundation, 1959.

Carnegie Corporation of New York, "The Gross Educational Product: How Much Are Students Learning?", Carnegie Quarterly, XIV, No. 2, 1966.

Cartter, Allan M., An Assessment of Quality in Graduate Education. Washington, D. C.: American Council on Education, 1966.

Chronback, Lee J., "Course Improvement Through Evaluation." Teachers College Record, May 1963, pp 672-683.

Committee of the Secondary School Examinations Council Appointed by the President of the Board of Education in 1941,(Great Britain), Curriculum and Examination in Secondary Schools. London: His Majesty's Stationery Office, 1943.

Dressel, Paul L., et al., Evaluation in Higher Education. Boston: Houghton Mifflin Company, 1961.

Dressel, Paul L., and Mayhew, Lewis B., General Education: Explorations in Evaluation. Washington, D. C.: American Council on Education, 1961.

Eckert, Ruth E., Outcomes of General Education. Minneapolis: University of Minnesota, 1943.

Educational Testing Service, A Plan for Evaluating the Quality of Educational Programs in Pennsylvania-Highlights. Princeton: Educational Testing Service, 1965.

Flanagan, John C., et al., Design for a Study of American Youth. Boston: Houghton Mifflin Company, 1962.

Furst, Edward J., Constructing Evaluation Instruments. New York: Longmans, Green and Co., 1958.

Haggerty, Melvin E., The Faculty, Vol. II of The Evaluation of Higher Institutions. Chicago: The University of Chicago Press, 1937.

Hyman, Herbert H., et al., Applications of Methods of Evaluation: Four Studies of the Encampment for Citizenship. Berkeley and Los Angeles: University of California Press, 1962.

Keppel, Francis, The Necessary Revolution in American Education. New York: Harper & Row, Publishers, 1965.

Klineberg, Otto, et al., Evaluation Techniques - International Social Science Bulletin. UNESCO Vol. VII, No. 3, 1955.

Lindquist, E. F., ed., et al., Educational Measurement. Washington, D. C.: American Council on Education, 1951.

Merwin, Jack C., and Tyler, Ralph W., "What the Assessment of Education Will Ask", Nation's Schools, November, 1966. pp 77-79.

National Advisory Mental Health Council, Evaluation in Mental Health. Washington, D. C.: Public Health Service, 1955.

National Education Association, Profiles of Excellence. Washington, D. C.: National Education Association, 1966.

National Study of Secondary School Evaluation, Evaluative Criteria. 1960 Edition, Washington, D. C.: National Study of Secondary School Evaluation, 1960.

National Study of Secondary School Evaluation, Evaluative Criteria for Junior High Schools. Washington, D. C.: National Study of Secondary School Evaluation, 1960.

Remmers, H. H. and Gage, N. L., Educational Measurement and Evaluation. New York: Harper & Brothers, 1943.

Secondary School Examinations Council, Committee on Curriculum, Curriculum and Examinations in Secondary Schools (Great Britain). London: His Majesty's Stationery Office, 1943.

Seiler, John A., Systems Analysis in Organizational Behavior. Homewood, Ill.: Richard D. Irwin, Inc. and The Dorsey Press, 1967.

Thorndike, Robert L. and Hagen, Elizabeth, Measurement and Evaluation in Psychology and Education. New York: John Wiley & Sons, Inc., 1961.

Tucker, Ledyard R., "Formal Models for a Central Prediction System." Psychometric Monographs, No. 10, 1963,

Tyler, Ralph W., "Assessing the Progress of Education in Science", The Science Teacher, September, 1966, pp 11-14.

Tyler, Ralph W., et al., Perspectives of Curriculum Evaluation. Chicago: Rand McNally & Company, 1967.

D

4

D. APPENDICES

4. Case Study: The Application of the Bressler-Tumin Model of Educational Goals and Outcomes to Materials Published by the New Jersey Department of Education and the Princeton, New Jersey Regional Schools

Iseli K. Krauss

and

Margaret M. Waldron

Case Study

The Application of the Bressler-Tumin Model of Educational Goals and Outcomes

to

Materials Published by the New Jersey Department of Education
and the Princeton, New Jersey, Regional Schools

Iseli K. Krauss
Margaret M. Waldron

In order to test the usefulness of the Bressler-Tumin models of educational goals and outcomes[1] in actual school systems, a small case study was undertaken. Officials of the New Jersey Department of Education in Trenton and of the Princeton (New Jersey) Regional Schools were contacted and their cooperation obtained. They made available, in discussions and through published materials, a great deal of information regarding the functions and achievements of the offices and schools.

In discussions, and while studying materials provided by the schools, it became apparent that there were enormous gaps in the knowledge the State's Department of Education had about actual goals and accomplishments of the local schools. It also quickly became evident that the local schools had a somewhat clearer idea of their own accomplishments than the Department of Education had of the local accomplishments. There were great differences between the amount of information the local schools had and that requested or collected by the State; the local schools had much more information on accomplishments than the State required.

It seemed important to describe briefly the kinds of information the schools have of their own goals and accomplishments, the information the Department of Education has concerning the goals and accomplishments of the local schools, the discrepancies between the State and local information, and the probable reasons for these discrepancies. A description of the case study will be found in Appendix A.[2]

---

[1] See Appendix B for an outline of the model used.

---

[2] The case study was carried out over a period of several months by Iseli K. Krauss and Margaret M. Waldron, research assistants.

THE NEW JERSEY DEPARTMENT OF EDUCATION

The Department of Education is composed of several divisions under the overall supervision of the Commissioner of Education. Within the Division of Administration are Personnel Services, Public Information, Print Shop, Teacher Education and Certification, Adult Education, Academic Credentials, Civil Defense, Adult Education Project, Machine Services, and other miscellaneous offices. Two offices of major concern to the investigation, Office of Statistical Services and Federal Assistance Programs, are also in this division.

The Division of Business and Finance covers matters of budget, school planning including architecture, building construction and transportation. Controversies and Disputes are handled in a division set up for that purpose. The State Library, Archives and History constitute one division, The State Museum, another.

The Division of Curriculum and Instruction is quite large with sections for Secondary Education, Primary Education, Guidance and Testing, Special Education, Early Childhood Education and others. In addition, the Division makes use of consultants in several fields such as arts and humanities, science, English and reading. The Division of Vocational Education is also quite large , covering areas such as private trade shcools, veterans training, agricultural education, home economics, technical training, industrial arts, office occupations and business education. County coordinators for vocational education (Camden County, Gloucester County and Hudson County) and the Newark Manpower-Training

Skills Center are also covered by this division. The Marie H. Katzenbach School for the Deaf is in a division of its own.

Five offices were recommended as being the most likely to have goal and outcome material available.

(1) Division of Business and Finance

(2) Division of Vocational Education

(3) Office of Statistical Services (Division of Administration)

(4) Federal Assistance Programs (Division of Administration)

(5) Division of Curriculum and Instruction

A description of the materials each of these divisions provided and the type of information contained in each publication follows.

(1) Division of Business and Finance

The primary contribution of this office was the Fifteenth Annual Report of the Commissioner of Education, a summary of financial statistics of school districts. Revenues, expenditures, enrollment and personnel data for each school district in the state are presented in a format allotting one page per school district. Summary pages for counties and for the entire state precede the 595 local district reports. No mention of goals or outcomes of education is made. The book merely states where the money for each district comes from and how it is spent. In theory it might be possible to guess what aspects of the educational program are most important to a particular district by studying how the available money is spent and from that information determine the educational goals of a district. But that seems risky. One must assume that the money spent reflects the importance of a program, and furtherthat one can know all the possible educational goals from which the local administration had chosen the ones they would try to achieve. The summaries of expenditures in the Annual Report are too general for these assumptions to be met.

It was in the offices of the Division of Business and Finance that a copy of Division Accomplishments of 1965-1966 and Goals for 1966-1967 came to light. Although no one had suggested that it might be relevant, it proved to be one of the most productive sources of goal and outcome statements. Each year, all divisions submit statements of accomplishments for the preceding year and separate statements of goals for the forthcoming year. The accomplishments are listed in the first section of the book and the goals in the latter section. White it is true that more statements of educational aims and educational achievements were gathered from this book than from any other single source, the total number of separate goal statements found was smaller than might have been expected, given the large number of officials and divisions reporting. The goals and accomplishments listed were not so much those of education as of administration. The Division of Curriculum and Instruction listed all the conferences attended by representatives of the Division, the Center for Psycho-Social Studies listed meetings attended on such matters as research design and methods of recording student data. These Divisions, then, outlined accomplishments and goals of their own personnel and offices rather than those of education *per se* in the State.

In this particular report, the emphasis is on formal efforts made from these offices to improve education in some way or to change local school conditions -with little evidence presented of impact or effects of these efforts.

All this is not to say that the goals and outcomes of the divisions of the State Department of Education are irrelevant to education. The goals described for the divisions in most cases are intended to make better education available to more students whatever their district. The stated goal of the Division of Curriculum and Instruction, "To work with the problems of assessment as a key to instructional improvement, " while not a goal of education in terms of desired behavioral characteristics of students, can certainly be thought of as potentially contributing to certain desired behavioral characteristics of students.

(2) Division of Vocational Education

Most of the data for this division was obtained from the record of division accomplishments and goals discussed above. This division stated its accomplishments and goals in a very different fashion from the other divisions. Many accomplishments and goals were in terms of specific behavior hoped for or observed for vocational students in each of many areas. For example, there was specific mention of programs intended to provide potential drop-outs with vocational skills, programs intended to help nurses who had not been active in the profession recently to regain nursing skills and programs to provide jobs for vocational program graduates. This particular division originates and directs many programs. Most of the other divisions tend to act in an advisory capacity with their aims being to aid and assist the local school districts achieve their aims.

(3) Office of Statistical Services

The director of the Office of Statistical Services provided a comprehensive understanding of the central record-keeping functions of the Department of Education. Information on finances, enrollment, etc. was readily available, but substantive information on the outcomes of education for New Jersey students was not being handled by this Office. Much information on educational outcomes was collected in a rather random fashion by various offices, but it was not on a form that the Statistical Services staff could use, had it been collected by a single office.

Several interesting and comprehensive publications are issued by this office. One with a long self-explanatory title is: Survey of Grade Span, Enrollment, Pupils on Half Session, Pupils in Substandard Classrooms, and Number of Certificated Personnel in Each New Jersey Public School by County and District. One of the most interesting publications is the summary report of teacher vacancies circulated each August and January. This report lists all vacancies for each school in the state, making schools aware of vacancies in other schools. Referrals of candidates are made simply and quickly on the basis of these reports. This publication does not deal directly with educational goals or outcomes but is certainly not of negligable importance to the education process.

(4) Federal Assistance Programs

The Title I programs funded by the Elementary and Secondary Education Act of 1965 are coordinated by members of the New Jersey Department of Education. Each local school district wishing to receive funds for a project under Title I must submit a proposal to the Title I Office for approval. At the completion of the project, an evaluation is carried out by the state Title I Office based on evidence submitted by the local schools and on a visit to the school itself by a Title I official. The proposals and evaluations must include goals and outcomes. The state office then compiles tables of projects completed with the aims and outcomes of each. The schools must choose among objectives suggested on the Title I application form and must report accomplishments by choosing among alternatives listed on the evaluation form. Spaces for "Other" are provided but the information in that category from one school is put together with information from a different school with a different program, and all together labeled "Other." The local schools choose among objectives such as: "To improve performance as measured by standardized achievement tests," "To improve children's verbal functioning," "To improve the children's self-image," "To reduce the rate and severity of disciplinary problems," and "To improve the physical health of the children." These objectives are among those listed for the five categories of objectives: Achievement, Ability, Attitudes, Behavior and Conditions Related to Learning. Each category includes an Other designation as well.

On the evaluation instrument, outcomes are designated by the following category names: Achievement, Attitudes, Behavior, Other, Teacher Growth and Parents. Sample objectives for these categories are respectively: "Reading," "Mental Health," "Citizenship," "Cultural Enrichment," ("Speech Correction" and "Physical Health and Fitness" also come under this Other category) "Professional Activity." and "Improved Home-School Relationships." In the evaluation process the schools also classify each project as Substantial Progress Achieved, Some Progress or Little or No Progress. Standardized test results are also reported if available.

Although the evaluations of the Title I projects are not carried out using the same descriptive terms used in the proposals for the projects, the State Title I office has on file large numbers of specific aims and outcomes for education projects. In the information collected on the Vocational Education Division there were many goal and outcome statements, probably because the programs were often originated or closely directed by the division. In contrast, the information on goals and outcomes of the Title I projects are available through the state offices because these projects must be approved and evaluated by the state Title I people.

(5) Division of Curriculum and Instruction

Elementary Education

Elementary education in the state of New Jersey is a very local matter. Elementary schools in this state do not need state approval of programs or curriculum but are charged with the responsibility to prepare students for high school. The only information the state offices have on elementary outcomes and objectives is informal and gathered during meetings with local school officials or at annual meetings, etc. Some Title I information could be expected to get back to these offices, too. The function of the Elementary Education office is almost entirely a consulting one. Information and assistance are available to the schools whenever requested.

A committee has recently been established to develop a guide book for self-evaluation at the classroom level for elementary schools. Though there will be no state requirement to complete such self-evaluations, there are great hopes that many schools will take advantage of the opportunity for self-evaluation.

The director of this office is quite enthusiastic about the prospects of the self-evaluation but, for a variety of reasons, not in favor of state-wide evaluation of elementary schools on the basis of standardized test results. The major objection is that it is unfair to compare test scores of students from schools serving communities with differing needs and educational programs.

Two of the publications of the elementary education office were concerned with educational practices, one reporting the various practices employed in the state during the school year 1962-1963 and one reporting practices discusse at a conference on educational innovations being carried out by New Jersey elementary schools. This office also made available a pamphlet, "This We Believe," consisting of a wide ranging statement of educational objectives. This pamphlet was not published by the Department of Education, but by the Department of Elementary School Principals of the New Jersey Education Association.

Secondary Education

The Department of Education has a much more direct role in the functioning of local high schools than of elementary schools. Secondary schools must be evaluated every five years to obtain state approval. This approval which makes the diploma valid and permits the transfer of credits is attendant upon the following:

a. A self-study carried out by the faculty and staff of each secondary school

b. The submission of statistics

c. A visitation by the County Superintendent and a representative of the New Jersey Department of Education

a. Self Study

For the required self-study, a booklet entitled Guidelines for Approval Through Self-Study for New Jersey Secondary Schools is furnished each school. Certain parts of the School Law are reported along with several requirements of the State Board of Education. Of the 14 School Law requirements presented, 10 concern physical health, mental health and safety; two, patriotism; one, United States history and one, teacher certification. Of the 14 State Board of Education requirements presented, four concern the granting of diplomas and academic credits; five, health and safety; one, administration time responsibilities; two, athletic contests and two, special classes for handicapped children.

An outline for comprehensive study is also suggested in this booklet and includes the following topics:

(1) Administration
(2) Supervision and Curriculum Development
(3) Instruction
(4) Instructional Materials
(5) Pupil Personnel Services
(6) Co-curricular Program
(7) Health and Nutrition Services
(8) Community Relation
(9) School Plant, Site and Equipment

b. Submission of statistics

The required statistics report certain educational outcomes as well as administration and teaching information. The following data must be submitted:

Enrollment figures

Distribution of classes by size

Student class load

Letter grade information

Graduate follow-up information

Academic offerings

Activities and clubs

Teaching load

Professional preparation of staff

Staff certification

Staff membership

Staff experience and length of service

c. Visitation

The County Superintendent of Schools visits the school either by himself or in company with members of his staff for the purpose of reviewing the self-study with the local administration and faculty. The visitation affords an opportunity to discuss the problems of the school, review and suggest plans for the future and to suggest possible sources of aid and assistance.

In addition to the approval by the State Board of Education, many schools seek approval every ten years by the Middle States' Association of Colleges and Secondary Schools.

The Office of Secondary Education published a pamphlet entitled "High School Organization in New Jersey" which offers a general introduction to the schools in the state. General purposes in the broadest of terms are related. The statements are phrased in such a way that it is quite clear that not every school embraces all of these aims. One such cautious statement is: "The better public high schools endeavor to provide for most of the educational needs of all the youth of the community." Statewide outcome data appears occasionally in the periodical "Secondary School Bulletin." The February 1968 issue was devoted to the follow-up of the New Jersey public high school graduates of the previous June.

Additional goal and outcome information for this Division was found in the Division Accomplishments and Goals discussed on page 4.

Limitations on Goal and Outcome Data Collection

Nearly everyone contacted stated with no hesitation that more and better evidence of educational outcomes was needed. Several of the state representatives had page after page and book after book of statistical information, none of which had any direct evidence of educational outcomes. There were statistics covering such aspects of the education system as school-by-school finances, architectural requirements and the number of conventions attended. None of this information is unimportant. It does not, however, give direct evidence of educational goals or outcomes.

Some specific limitations on the usefulness of the information gathered are listed below.

(1) The administrative nature of the functions of the New Jersey Department of Education limits the amount and type of goal and outcome information offered or collected by the Department. The publications of the Department are primarily activity reports of the various Offices and give little evidence of the effect on educational outcomes of these activities.

(2) The sources of goal and outcome material reported by the State are not identical. More goal statements originate from the Department of Education than from local schools; outcome statements refer to and depend upon information provided by the local schools rather than by the Department of Education.

(3) The schools of New Jersey are primarily under local control. Much evaluation information and policy information remain with the local school.

(4) There is little use made of standardized test results outside the local schools except for fulfilling the Title I program requirements.

(5) Although every Office contacted in the Department of Education was cooperative and willing to supply whatever materials they had, it was difficult to obtain relevant information. For instance, a good deal of effort was spent in trying to locate an important study on the New Jersey high schools which was prepared for the State and an independent outside organization. Both the

independent organization and the Department of Education referred us to the other group as neither could find a copy of the written report, although one copy of the statistical report of the study was eventually found.

THE PRINCETON, NEW JERSEY, REGIONAL SCHOOLS

Princeton public schools are organized on a 5-3-4 basis. That is, the elementary schools comprise grades kindergarden through five; the middle schools, grades six through eight; and the high school, grades nine through twelve. The materials from the regional schools were collected from each school of the system and transmitted to us by the Assistant to the Superintendant, over a period of several weeks.

Regional Information

A statement of School Board Policy for the Princeton Regional Schools was issued in 1966 when the Township and Borough schools were consolidated. The statement, actually a series of many statements, covers matters of policy from class size to faculty leaves of absence. The first statement, a general statement of philosophy and purposes, was originally used by the Princeton Township schools (it is dated 1964) but apparently not by the Borough Schools. According to school officials, the statement can be assumed to represent regional policy. Nine listed concerns of the schools take into consideration academic competencies, health, ethical behavior and standards, leadership, esthetics and social and emotional development.

The primary means of communication between the Board of Education and the district residents is an official newsletter of the Princeton Regional Schools, Regional Profile. News of educational programs and other school information of general concern is mailed to all district residents several times a year. Space restrictions permit only general discussions of most programs and issues, but a lead article typically deals with one particular aspect of the schools in detail. Both aims and outcomes are offered in the lead article and in the briefer articles.

The New Jersey Department of Education requires volumes of statistical information each month. Some idea of this information was obtained during the examination of the state information. But the full impact of the amount of data compiled each month was not felt until the local information was studied. The Department of Education issues a mimeographed booklet with a month-by-month list of required reports. In some months, there are nearly 20 reports to be made. The information reported covers such topics as school milk programs and absences, among others.

The Annual Report of Educational Statistics is due only once a year, however. The reports submitted by the local districts to the state offices are extensive, with breakdowns of many figures. Subsequently, each district's data are condensed to one page by the state offices for presentation in the Annual Report of the Commissioner of Education. This volume and the information contained in it are discussed on page 4 in the section on the Division of Business and Finance.

Elementary Information and Middle School Information

Information on the goals of education of the elementary schools of the Princeton Regional Schools was difficult to find. Each school issues a yearly handbook with the cooperation of the Parent Teachers Association that is primarily concerned with procedural information. Such issues as school closing for snow, acceptable dress and birthday parties are covered. The inevitable section on the purposes of homework often contains goal statements. The booklet for the Middle School does present several aims of education. One such is the following: "Students will be encouraged to develop their talents to their full potential in all areas from academic to activity interest."

In a statement to the Board of Education in April, 1968 by the John Witherspoon School, innovative educational programs for the forthcoming year were proposed. After a discussion of the rationale for the programs, specific plans for team teaching, multi-age grouping and non-grading were outlined. Each of these proposed innovations was discussed in terms of possibilities for improved teaching and increased benefit to the students involved in the programs.

Outcome statements are available in the form of scores on standardized tests such as the "Iowa Tests of Basic Skills". Test scores are reported for language skills, work-study skills and arithmetic skills with further breakdowns. Building averages and individual scores are reported. These scores are kept on file by the local schools and are not forwarded to the state offices.

High School Information

The administration of Princeton High School issues a booklet presenting the curriculum of the school with a brief description of every course offering. Both educational goals and outcomes are discussed in this booklet which was a major source of statements for the case study.

In a recent report to the Board of Education, the vocational guidance staff of Princeton High School reviewed the ways by which they are able to help students, primarily non-college bound students, reach vocational and personal adjustment. The goals which the staff hoped the students would reach were listed and discussed.

Outcomes can be extracted from the publication outlining preliminary plans for the graduating class of each year and the supplementary follow-up of those ex-students employed.

The Princeton High School self-study section of the five-year evaluation was made available. The present strengths (1967-1968), the needs, and the plans to remedy the needs were listed. Very few of the topics covered in the self-study discussed aims or outcomes of the students. They dealt, rather, with the procedures to be followed by the school administration and faculty in their efforts to improve conditions in the school and to improve the education of the students in the system.

Academic letter grades for the High School could have been available at a low cost. The individual grades for each student are on tape and could be compiled by computer into a composite record of the school. This particular composite information is used rarely by the school. It is required, however, on the forms of the five-year evaluation.

Summary of Local Information

Existing policy statements frequently deal with administrative and school characteristics rather than with desired educational outcomes of the Princeton Regional School students. Plans for new programs and teaching methods may be discussed in detail in evaluation reports but give no indication how these programs are hoped to influence the children. One must assume that "freeing the teacher from unnecessary paper work" will automatically be of value to the children but the nature of the value is not expressed. The proposal for innovative programs in the John Witherspoon School discussed on page 18 provides a notable exception to this type of report.

Outcome material does exist for the Princeton schools. Each child's record is extensive, but composite records other than classroom Iowa Test Scores are not kept. One exception is the record of letter grades submitted to the state each five years by the High School.

In general, the High School has more goal and outcome material than do the elementary or middle schools. This is mainly due to the state requirements for a five-year evaluation. As the elementary schools have a very high degree of autonomy in comparison to the local high schools, and are not required by the state to justify curricula, they may not find it necessary or useful to set forth goals and outcomes in a systematic way.

APPENDIX A

For the case study of the usefulness of the Bressler-Tumin model of educational outcomes in the assessment of school achievements and the classification of educational goals, both the New Jersey State Department of Education and the Princeton (New Jersey) Board of Education were contacted. It was felt that data from both state and local sources would provide a better idea of the kinds of information available than data from one school system alone or even from several local school systems. The primary goals of the case study were to determine, as follows,

1) what kinds of evidence of school achievement the State required or desired,
2) what kinds of evidence local schools actually gave to the State to determine what sorts of statements of purpose or desired outcomes the State Department of Education and the local schools had formulated and what use, if any was made of these statements, and
3) to extract from all our collected data whatever statements of goals and educational outcomes that could be found, record them, and classify them according to the Bressler-Tumin model.

The case study was begun by corresponding with the Commissioner of Education of New Jersey who cooperated very generously by making it possible to confer with several officials. The investigators corresponded with, or met, the following Department of Education representatives:

Carl L Marburger, Commissioner of Education

Robert S. Fleming, Assistant Commissioner of Education,
Director, Division of Curriculum and Instruction

William H. Warner, Director of Secondary Education,
Division of Curriculum and Instruction

Anne S. Hoppock, Director of Elementary Education,
Division of Curriculum and Instruction

Edward W. Kilpatrick, III, Assistant Commissioner of Education,
Division of Business and Finance

Harold Y. Bills, Director, Bureau of Business Services, Education,
Division of Business and Finance

S. David Winans, Director, Office of Statistical Services,
Division of Administration

Louis A. Dughi, Coordinator, Federal Assistance Programs,
Division of Administration

Gustav H. Ruh, Title I Coordinator, (Elementary and Secondary Education Act)
Division of Administration

Ira Helfgott, (Title I Elementary and Secondary Education Act)
Division of Administration

W. H. Rhodes, Assistant to the Superintendent of the Princeton Regional Schools, contributed substantially to our understanding of the Princeton school system during later meetings.

Each of these people contributed whatever information he had available that could possibly be relevant to the study. (A list of the publications and related items received and studied will be found in Appendix C.) As the papers were examined for statements of purpose or "goal statements", it was necessary to distinguish between clear statements of purpose and statements that implied purpose. For the purposes of the study, a goal statement was defined as a statement indicating some desired behavior of a student (or students) during his years in Kindergarden through Grade 12 or in his later life as a result of his education. That is, a goal of education could be thought of as (a) the desired behavior or trait of a student or students in school or later that (b) would come about or be increased by the educational process. Such a statement of goals is illustrated by the following:

> "All youth need to understand the methods of science, the influence of science on human life, and the main scientific facts concerning the nature of the world and of man."

Many statements encountered, however, did not meet the standards of the working definition. They were often called goals by their creators, but dealt only indirectly with students. The main focus of such statements was on action to be taken by the school or, often, by teachers, and not on any anticipated student behavior. These statements implying goals as defined by the study were, then, those making no direct statement of desired outcome, but nevertheless suggesting strongly that some behavior or trait was considered to be desirable and the likelihood of its occurrence would be increased as a result of a particular administrative action or of a particular application of the education process. A statement illustrating an implied goal is the following:

> "It has been recognized that the elementary student's learning process could be developed more rapidly through classroom instruction involving the use of materials and material processing."

In order to use as many of the statements implying educational goals as possible, several were reworded into a form meeting the requirements of the operational definition of a goal statement. The reworded form of the above statement is:

> "It is important that the student's learning process be developed more rapidly."

Having made the distinction between clear statements of educational goals and statements implying educational goals, a similar distinction was also found to be useful in looking at claimed outcomes, though to a lesser degree. A claimed outcome was defined as a statement of observed behavior or trait of a student or ex-student that presumably came about as a result of the education process. Such a statement is the following:

> "All vocational graduates were placed in good positions before school closed in June."

A statement implying a claimed outcome is illustrated by the following:

> "Multi-occupation skill centers were established in Newark, Camden, Salem County and Trenton to provide vocational training and basic education for potential workers who lack the skill and education needed to compete for positions in the labor market."

Each statement implying an outcome was recorded as it was originally written and then a statement describing the outcome was added in a form consistent with the more straightforward statements. The implied outcome of the above statement became:

> "Potential workers learned the skills and obtained the education necessary to compete for positions."

Outcomes supported by evidence and those with no supporting evidence were not differentiated. A claim such as "Our students are prepared to take on the responsibilities of citizenship" was given equal value as a statement claiming that "74% of the graduates of our school went on to further education." If an outcome regarding behavior of students was claimed, it was included whether or not it could be substantiated. (Had only claims with supporting evidence been included, there would have been few claims reported.)

There were now four types of statements to classify:

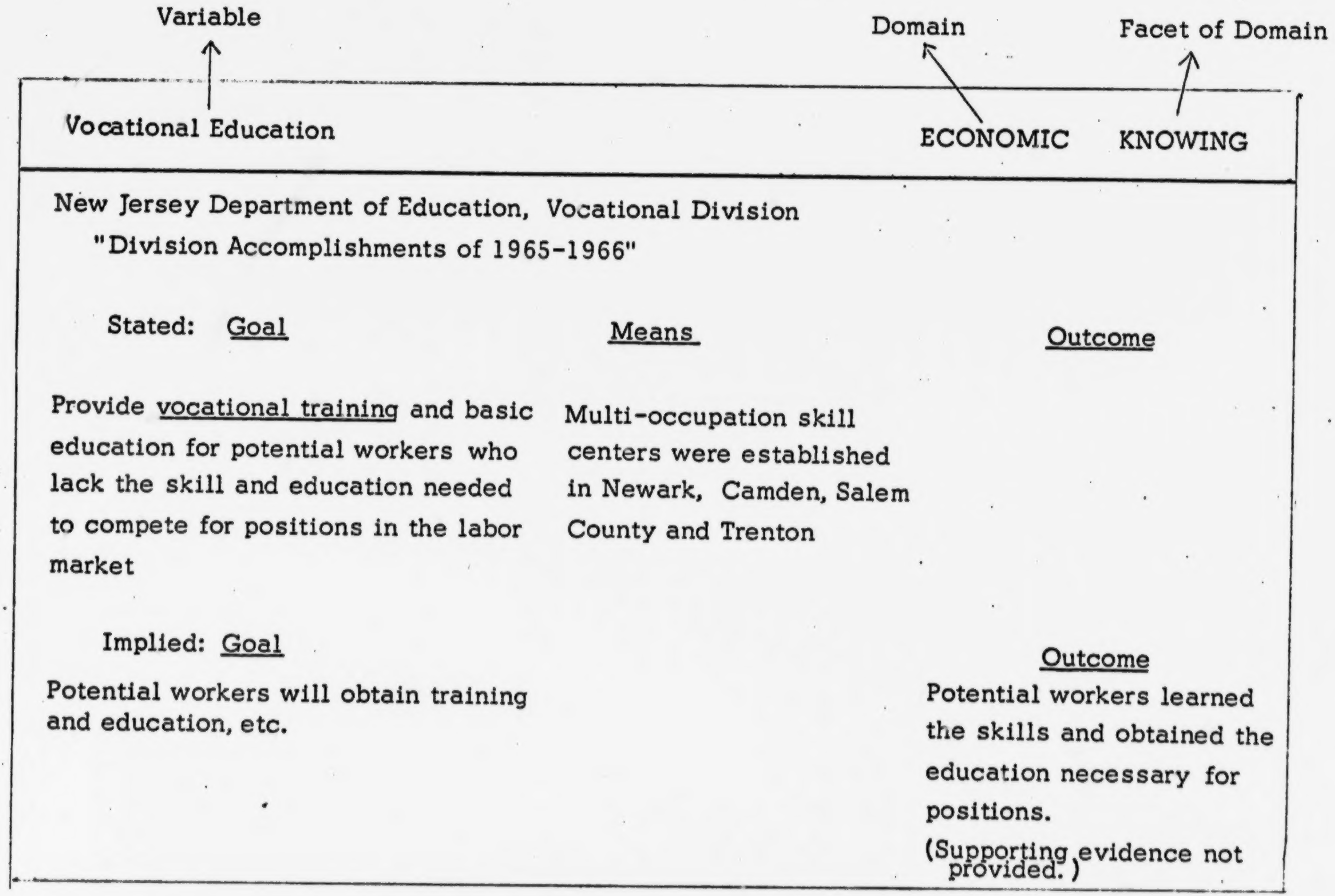
(1) Clear statements of purpose
(2) Statements that implied purpose, rewritten into clearer statements of purpose
(3) Clear statements of claimed outcomes
(4) Statements that implied outcomes, rewritten into statements of claimed outcomes

The statements were recorded on colored five-by-eight-inch index cards. Green cards were used for goals listed by the Department of Education, yellow cards for outcomes of education claimed by the Department of Education, blue cards for goal statements of the Princeton Regional Schools, and pink cards for claimed outcomes of the Princeton Regional Schools. For reference purposes, each statement was given a "variable name" to identify the area of information contained in the statement in as few words as possible. In some cases, the variable name chosen represented the goal, in some, the particular method employed, and in other cases, the outcome. Examples of the variable names included are: mathematics, test scores, occupation. Admittedly the labeling was inconsistent. In each instance, the card was labeled according to the primary emphasis of the statement in question. Had time been unlimited, separate cards for each part of the statements would have been made for the goal, the method, the outcome, or for the particular students involved -- for whatever was relevant. (See Figure 1 for a sample goal card and Figure 2 for a sample outcome card.)

If more than one goal or outcome were contained in a statement, separate cards were made for each goal or outcome and the particular section of the statement being classified on that card was underlined.

Brief summaries of the documents and the purposes of these documents were recorded on white cards and sorted according to state or local source. The file of these cards provided an annotated bibliography of the documents studied.

Once the statements and the summaries of the documents were recorded, the classification was begun. Only two areas of the Bressler-Tumin classification scheme were used, Domains and Facets; and of the latter, only the three major sections were used. That is, each statement was classified as being predominately concerned with one or more of the


Variable
Domain
Facet of Domain

Vocational Education

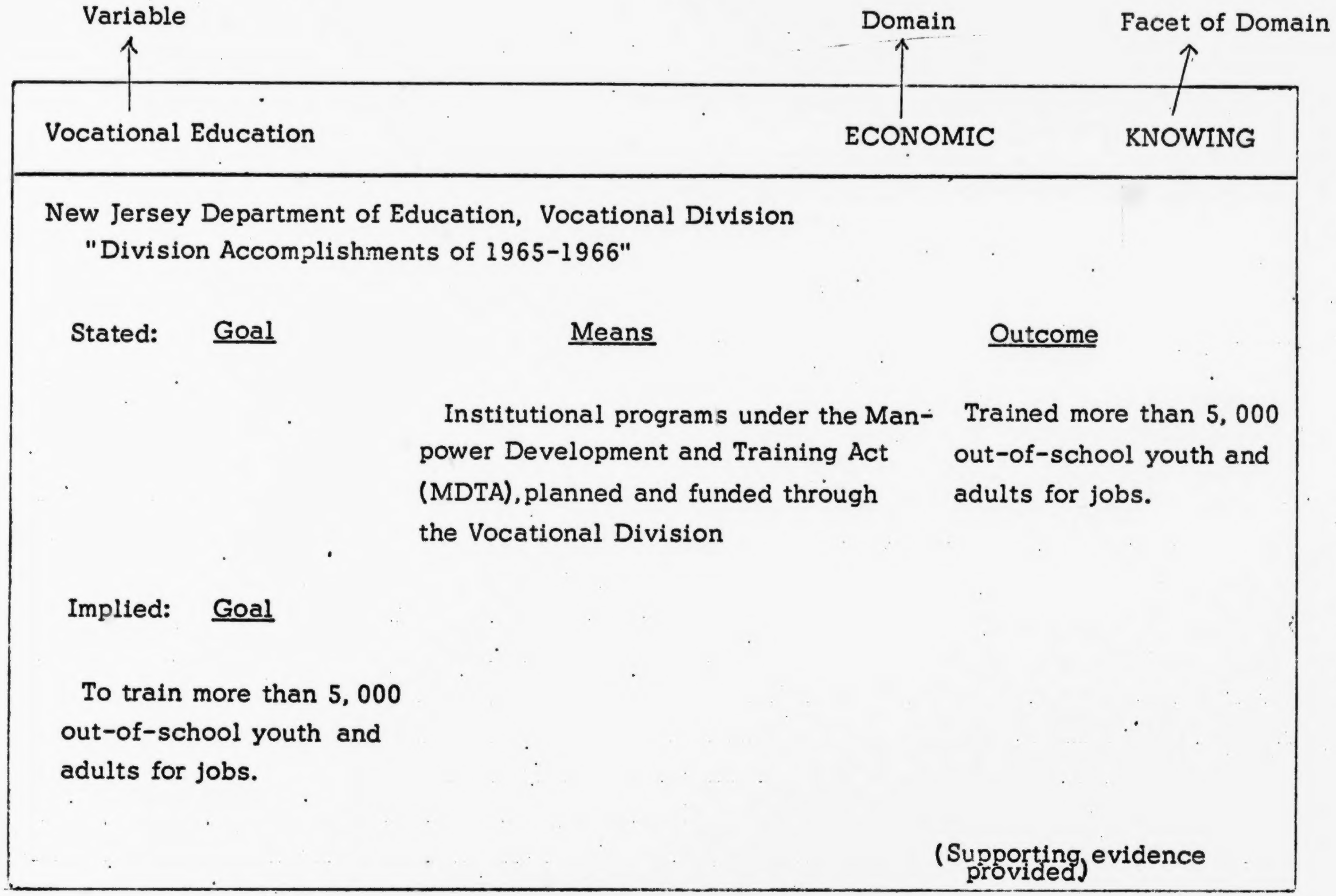
ECONOMIC KNOWING

New Jersey Department of Education, Vocational Division
"Division Accomplishments of 1965-1966"

Stated: Goal	Means	Outcome
Provide vocational training and basic education for potential workers who lack the skill and education needed to compete for positions in the labor market	Multi-occupation skill centers were established in Newark, Camden, Salem County and Trenton	

Implied: Goal		Outcome
Potential workers will obtain training and education, etc.		Potential workers learned the skills and obtained the education necessary for positions. (Supporting evidence not provided.)

Figure 1
Sample Goal Card



Vocational Education ECONOMIC KNOWING

New Jersey Department of Education, Vocational Division
"Division Accomplishments of 1965-1966"

Stated: Goal	Means	Outcome
	Institutional programs under the Manpower Development and Training Act (MDTA), planned and funded through the Vocational Division	Trained more than 5, 000 out-of-school youth and adults for jobs.

Implied: Goal

To train more than 5, 000 out-of-school youth and adults for jobs.

(Supporting evidence provided)

Figure 2

Sample Outcome Card

seven Domains and one or more categories of the "knowing-valuing-doing" dimension. If a statement clearly fell within more than one category of one or both the dimensions, duplicate cards were made for each classification applicable. If it seemed impossible to classify a statement in one or the other division, it was classified in one and listed as unclassifiable in the other. Statements of this type usually were worded so broadly that all the categories of one division applied. An example: "Forty-two Title I projects achieved substantial progress in the area of cultural enrichment." Not enough information was given to make a judgment on the knowing - valuing - doing dimension.

All the cards were classified by both researchers, working together with the first cards and working independently with the latter cards. There was substantial agreement on the classification in almost every case. However well or poorly the statements were being classified, they were being classified in the same way.

Difficulties with the scheme

Many difficulties were encountered in classifying the statements by the scheme dimensions. Most of these difficulties were also found by A. Smith and J. Smith in their earlier application of the scheme to statements of educational outcomes found in the literature. Rather than reiterate all the problems met, a few of the major problems will indicate possible inadequacies of the scheme. Basically, there was a great overlap among the categories and some statements did not readily fit into any of the categories. (Most of the difficult-to-classify statements were finally assigned to the Domain categories of Cultural Heritage, if of an academic nature, or to Orientations, if of a personal nature.) In addition, the categories did not always cross-cut. For example, there were several instances of "knowing knowledge."

Though it was very easy to construct statements that clearly fell within one category of Domain or Facet, it was much more difficult to fit statements written by others into our categories. Just such a simple matter as the particular phrasing of a statement by a school official could make classification difficult, not always because the statement itself was a poorly written statement of purpose or outcome, but because the specific wording made it imperative to classify the statement one way while the intent of the statement might have been classified quite differently. For instance, the statement, "The student learns to type business letters" was classified as a "knowing" statement - (the student knows how to type business letters) rather than as a "doing" statement - (the student types business letters). Probably, the student in question actually did type business letters but that called for an assumption the investigators could not justify. All such statements of learning were classified as "knowing" statements with the justification that learning how to do something does not necessarily mean that the something is actually done. Statements of this nature differed from statements implying purpose in that the former clearly stated purpose, although written so that they were classified in a way that might not have agreed with the intent of the statement, while the latter merely implied purpose.

Other difficulties were handled in a similarly arbitrary manner. Within each classification problem area, however, statements were assigned to categories consistently. That is, each statement reflecting a classification problem, such as that illustrated above, was classified the same way as the others reflecting the same problem.

Summaries of Classification Problems

The findings listed below are of the most superficial nature. The numbers in most cases are small, the classification of statements was carried out by only two people, no statistical measures were taken, and as indicated above, problems existed with the scheme. Yet, certain of the tentative findings are interesting, even fascinating, and certainly worth looking at. The four charts on the pages following - (Figures 3, 4, 5, and 6), represent the application of the model to statements of local (Princeton) goals, local outcomes, state (New Jersey) goals and state outcomes. The findings below are taken from these charts. Many more could be pulled from the charts or from a comparison of the charts. The ones selected are but a beginning.

(1) The cell combining Cultural Heritage and Doing had the greatest frequencies for goals and outcomes, local and state.

(2) Of the Knowing-Valuing-Doing categories, the Doing category had the greatest frequency of entries. Many of the statements classified as Knowing might have been classified as Doing, had the wording of the statements been slightly different. (See the example on page 28.) Had that been the case, the frequency of classification for the Doing category would have been higher than it was, and the frequency of classification for the Knowing category correspondingly lower. The chance wording of many statements, then, artificially lowered the frequency of classification in the Doing category and increased the frequency in the Knowing category. Of course, the bias could easily have been in the opposite direction, with statements classified as Doing rather than Knowing because of particular wording.

	Economic	Political	Primary Group	Socialization And/Or Education	Cultural Heritage	Orientations	Physical Well-Being	Not Classified	Totals
KNOWING	18	3	4	1	16	5			47
VALUING	2	2			5	5			14
DOING	17	2		1	29	19	4		72
NOT CLASSIFIED	1				9	1	3		14
TOTALS	38	7	4	2	59	30	7		147

Figure 3
Classification frequencies of goal statements found in New Jersey Department of Education Materials

STATE OUTCOME STATEMENTS

	Economic	Political	Primary Group	Socialization And/Or Education	Cultural Heritage	Orientations	Physical Well-Being	Not Classified	Totals
KNOWING	7	1			6				14
VALUING					1				1
DOING	11		1	7	28	9	1		57
NOT CLASSIFIED					13	3	3		19
TOTALS	18	1	1	7	48	12	4		91

Figure 4

Classification frequencies of outcome statements found in New Jersey Department of Education materials

LOCAL GOAL STATEMENTS

	Economic	Political	Primary Group	Socialization And/Or Education	Cultural Heritage	Orientations	Physical Well-Being	Not Classified	Total
KNOWING		2	1		8	4	1	4	20
VALUING			1	1	1	2	1		6
DOING	3	2		1	29	12	4		51
NOT CLASSIFIED									
TOTALS	3	4	2	2	38	18	6	4	77

Figure 5

Classification frequencies of goal statements found in Princeton, New Jersey

Board of Education materials

LOCAL OUTCOME STATEMENTS

	Economic	Political	Primary Group	Socialization And/Or Education	Cultural Heritage	Orientations	Physical Well-Being	Not Classified	Total
KNOWING	4		1		8	1	1		15
VALUING					1				1
DOING	6			4	19	4	4		37
NOT CLASSIFIED									
TOTALS	10		1	4	28	5	5		53

Figure 6

Classification frequencies of outcome statements found in Princeton, New Jersey

Board of Education materials

(3) The greatest disparity between state and local frequencies was in the Economic Domain. There were 56 state entries and 13 local entries, probably because of the large number of goal and outcome statements provided by the Vocational Division of the Department of Education.

(4) Though very few entries were made for Valuing, more were made for state and local goals than for outcomes.

(5) The most frequent Domain of both goal and outcome statements was Cultural Heritage.

(6) Only one Political outcome entry was made and only a very few Political goal entries were made.

(7) There were 50% more goal entries made than outcome entries.

(8) There were twice as many empty cells for outcomes as for goals; the goals covered a much wider range of classification cells and therefore appear to be more varied than the outcomes.

(9) Primary Group had almost no entries. Possibly, elements that might have been classified here were classified under Orientations.

(10) The greatest difference between goals and outcomes was in Orientations, with a greater number of entries in goals. This difference might have been expected, for in this particular area it is very easy to make statements of aims but very difficult to point to outcomes.

Though many societies hope that education will alleviate the ills of the world as well as provide avenues for individual success, the New Jersey education officials and the Princeton Regional Schools hope for somewhat less, judging by their published statements. Most frequently, the pronouncements dealt with the area traditionally thought of as within the province of organized education -- reading, mathematics, social studies, etc. Furthermore, the officials claimed outcomes for the same areas for which they expressed aims. This can be partially, but only partially, explained by "companion" goals and outcome statements. An outcome statement was frequently accompanied by its own goal statement: "We hoped to send at least 70% of our seniors on to collge; actually 80% went on to some form of higher education." Outcome statements were more frequently paired with goal statements than goal statements with outcome statements for two reasons:

(1) there were more goal statements than outcome statements

(2) goals were stated for areas in which outcomes are most difficult to determine.

But even without these "paired" goal and outcome statements, there was considerable agreement among the classification areas of goal and outcome statements.

Appendix B

A. DOMAINS

Major Institutional Roles and Non-Role Structured Activities as Domains within Which Educational Goals are to be Located

1. Economic: including both production or service role and consumption role.
2. Political: including the varying forms of possible participation in the political process from holding office to membership in party to voting.
3. Primary Group Membership Roles: including both those connected to kinship and reproductive functions and structures, and those found in non-kin structures and relationships (e.g. friendship, love, etc.).
4. Socialization and/or Educational Activities: including everything concerned with the activities of the person as a socializer or educator of the young on the one hand, and his own involvement in education beyond the cut-off point here used to locate educational outputs.
5. Activities Connected with the Absorption, Understanding and Various Participations in the Cultural Heritage of the Society: including science, art, philosophy, religion, play, and recreation.
6. Orientations to Self and Others: including intrapersonal, interpersonal, and intergroup attitudes, where otherwise not classifiable as specific to another role area, and intended to refer to the generalized person.
7. Physical Well-Being: which we classify separately here, for no good reason other than that it is universally or nearly universally cited by every goal-stater and is not easily manageable under the above-listed categories.

C. FACETS OF DOMAINS

(Apply to each domain-aspect [Table B] in turn)

Types of Educational Goals or Aims of Outputs
Explicitly Affirmed or Desired, or Inferrable from, inter alia
Budgets of Money, Time, Personnel: Outcomes Evaluated: Honors Awarded:
Sources of Strain and Disappointment Recorded: Teacher+Characteristics
and Preparation Stressed: Curriculum Content Balance (Hours Allocated, etc.)

	KNOWING					VALUING			DOING		
COGNITION	RECOGNITION	DIVERGENT	CONVERGENT	EVALUATIONAL		APPROVING	PREFERRING		REPEATING	ELABORATING	CREATING

MATERIALS EXAMINED FOR GOAL AND OUTCOME STATEMENTS

New Jersey:

*SECONDARY SCHOOL BULLETIN, June 1967
Division of Curriculum and Instruction, N. J. Department of Education

*STUDY ON HUMANITIES
Center for the Humanities, Division of Curriculum and Instruction, N. J. Dept. of Education

*EVALUATION - TITLE I, PUBLIC LAW 89-10, FISCAL YEAR 1967
Office of Federal Assistance Programs, N. J. Department of Education

*HIGH SCHOOL ORGANIZATION IN NEW JERSEY
Office of Secondary Education, Division of Curriculum & Instruction, N. J. Dept. of Education

*SECONDARY SCHOOL BULLETIN, February 1968
Division of Curriculum and Instruction, N. J. Department of Education

*FINANCIAL STATISTICS OF SCHOOL DISTRICTS, SCHOOL YEAR 1965-1966,
(FIFTEENTH ANNUAL REPORT OF THE COMMISSIONER OF EDUCATION)
Division of Business and Finance, N. J. Department of Education

*DIVISION ACCOMPLISHMENTS OF 1965-1966 AND GOALS FOR 1966-1967
N. J. Department of Education

A REPORT ON EDUCATIONAL INNOVATIONS SURVEYED AT THE FAR HILLS INN CONFERENCE OF NEW JERSEY SUPERVISORS, October 21, 1966
Office of Elementary Education, Division of Curriculum and Instruction,
N. J. Department of Education

*THIS WE BELIEVE
Department of Elementary School Principals, New Jersey Education Association

*A STUDY OF PRACTICES IN THE ELEMENTARY SCHOOLS OF NEW JERSEY (1962-1963)
Office of Elementary Education, Division of Curriculum and Instruction,
N. J. Department of Education

*TITLE I PROJECT IDEAS
N. J. Department of Education

*GUIDELINES FOR APPROVAL THROUGH SELF-STUDY FOR NEW JERSEY SCHOOLS
Office of Secondary Education, Division of Curriculum and Instruction
N. J. Department of Education

FINANCIAL ACCOUNTING FOR NEW JERSEY SCHOOL DISTRICTS
Division of Business and Finance, N. J. Department of Education

SURVEY OF GRADE SPAN, ENROLLMENT, PUPILS ON HALF SESSION, PUPILS IN SUBSTANDARD CLASSROOMS, AND NUMBER OF CERTIFICATED PERSONNEL IN EACH NEW JERSEY PUBLIC SCHOOL BY COUNTY AND DISTRICT, September 1966
Office of Statistical Services, N. J. Department of Education

*Goal and/or Outcome Statements Extracted; the remaining publications did not contain goal or outcome statements

SUMMARY OF SELECT SCHOOL, CLASSROOM, PUPIL, AND PERSONNEL INFORMATION IN NEW JERSEY PUBLIC SCHOOL DISTRICTS, 1966-1967
Office of Statistical Services, N. J. Department of Education

INSTRUCTIONS FOR TITLE I 1968 APPLICATION FORMS
Office of Federal Assistance Programs; N. J. Department of Education

TEACHER VACANCIES, January 1968
Office of Statistical Services, N. J. Department of Education

INSTRUCTIONS FOR EVALUATION OF TITLE I FOR FISCAL YEAR 1967
Division of Curriculum and Instruction, N. J. Department of Education

MATERIALS EXAMINED FOR GOAL AND OUTCOME STATEMENTS

Princeton, New Jersey:

*JOHNSON PARK ELEMENTARY SCHOOL HANDBOOK FOR PARENTS
Johnson Park School Parent-Teacher Organization and Princeton (N. J.) Regional Schools

LITTLEBROOK ELEMENTARY SCHOOL HANDBOOK FOR PARENTS
Littlebrook Parent-Teacher's Organization and Princeton (N. J.) Regional Schools

*RIVERSIDE SCHOOL HANDBOOK FOR PARENTS
Riverside Parent-Teacher's Organization and Princeton (N. J.) Regional Schools

*JOHN WITHERSPOON ELEMENTARY SCHOOL HANDBOOK FOR PARENTS
John Witherspoon Parent-Teacher Association and Princeton (N. J.) Regional Schools

*REPORT OF AVERAGES WITH PERCENTILE RANKS OF GRADE 4 ON THE IOWA TESTS OF BASIC SKILLS (2/68 Midyear) - In John Witherspoon School
John Witherspoon School

*REPORT OF SYSTEM AVERAGES WITH PERCENTILE RANKS OF GRADE 4 ON THE IOWA TESTS OF BASIC SKILLS (2/68 Midyear) - All Princeton schools
Princeton Regional Schools

*PRINCETON MIDDLE SCHOOL PARENT HANDBOOK
Princeton (N. J.) Regional Schools

PRINCETON MIDDLE SCHOOL TEACHER HANDBOOK
Princeton (N. J.) Regional Schools

PRINCETON HIGH SCHOOL HANDBOOK
Princeton (N. J.) Regional Schools

*MEMORANDUM REGARDING GUIDANCE DEPARTMENT JOB DESCRIPTION
Princeton (N. J.) High School

*PROFILE - CLASS OF 1968
Guidance Department, Princeton (N. J.) High School

*GUIDANCE DEPARTMENT STATEMENT, SCHOOL BOARD MEETING, APRIL 1968
Princeton (N. J.) High School

*PRINCETON HIGH SCHOOL CURRICULUM
Princeton (N. J.) High School

PRINCETON HIGH SCHOOL TEACHERS' MANUAL
Princeton Regional Schools

*TEACHERS' MANUAL - ELEMENTARY SCHOOLS
Princeton (N. J.) Regional Schools

* Goal and/or outcome statements extracted; the remaining publications did not contain goal or outcome statements

*PRINCETON PUBLIC SCHOOLS' SCHOOL BOARD POLICY STATEMENTS
Princeton (N. J.) Regional Schools

*ANNUAL REPORT OF EDUCATIONAL STATISTICS FOR PRINCETON REGIONAL SCHOOLS FOR SCHOOL YEAR ENDING JUNE 30, 1967
Princeton (N. J.) Regional Schools

*PRINCETON HIGH SCHOOL EVALUATION REPORT
Princeton (N. J.)Regional Schools

COLLEGE GUIDE FOR JUNIORS
Guidance Department, Princeton High School, Princeton (N. J.) Regional Schools

*REGIONAL PROFILE
Princeton 'N .J.) Regional Schools

*SPECIAL PLANS FOR THE JOHN WITHERSPOON SCHOOL FOR THE 1968-69 SCHOOL YEAR
Princeton (N. J.) Regional Schools